

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 17:16:51 ; Search time 38.8905 Seconds

(without alignments)
8225.189 Million cell updates/sec

Title: US-09-902-481B-3

Perfect score: 5879

Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKROYKMMSEGGPGGABPQ 1137

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1155919 seqs, 281338677 residues

Total number of hits satisfying chosen parameters: 1155919

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5879	100.0	1137	10	US-09-902-481A-3
2	5857	99.6	1137	10	US-09-902-481A-4
3	5839	99.3	1153	9	US-09-350-259-3
4	5839	99.3	1153	10	US-09-902-481A-1
5	5839	99.3	1153	10	US-09-891-943-3
6	5839	99.3	1153	14	US-10-144-259-30
7	5839	99.3	1153	14	US-10-207-655-176
8	5836	99.3	1137	10	US-09-902-481A-5
9	5832	99.2	1137	10	US-09-902-481A-6
10	5823.5	99.1	1152	9	US-09-945-265-4
11	3446	58.6	1163	15	US-10-116-275-204
12	3423	58.2	1163	9	US-09-350-259-4
13	3423	58.2	1163	10	US-09-891-943-4
14	3388	57.6	1161	9	US-09-350-259-2
15	3388	57.6	1161	10	US-09-891-943-2

16	3372.5	57.4	1161	9	US-09-350-259-99
17	3372.5	57.4	1161	10	US-09-891-943-99
18	3213.5	54.7	1161	9	US-09-350-259-55
19	3213.5	54.7	1161	10	US-09-891-943-55
20	3201	54.4	1161	9	US-09-350-259-53
21	3201	54.4	1161	10	US-09-891-943-53
22	3192.5	54.3	1151	9	US-09-350-259-37
23	3192.5	54.3	1151	10	US-09-891-943-37
24	3180	54.1	1155	9	US-09-350-259-46
25	3180	54.1	1155	10	US-09-891-943-46
26	1848	31.4	369	12	US-10-087-192-1212
27	1536.5	26.1	1170	9	US-09-945-265-2
28	1532.5	26.1	1170	12	US-10-261-164-1
29	1509	25.7	1223	16	US-10-408-765A-295
30	1341.5	22.8	1086	16	US-10-408-765A-1871
31	1229.5	20.9	494	9	US-09-350-259-103
32	1229.5	20.9	494	10	US-09-891-943-103
33	1128.5	19.2	413	9	US-09-350-259-101
34	1128.5	19.2	413	10	US-09-891-943-101
35	1128	19.2	1179	15	US-10-177-550-2
36	1128	19.2	1179	15	US-10-173-551-2
37	1087	18.5	1188	15	US-10-291-265-810
38	1085.5	18.5	1151	10	US-09-984-130-103
39	1085.5	18.5	1151	10	US-09-836-353A-103
40	1085.5	18.5	1179	12	US-09-918-715-250
41	1083	18.4	1188	15	US-10-291-265-338
42	1079.5	18.4	1189	10	US-09-984-130-35
43	1079.5	18.4	1189	10	US-09-836-353A-35
44	1079.5	18.4	1189	12	US-10-262-839-4
45	1063	18.1	589	12	US-10-261-164-2

ALIGNMENTS

RESULT 1

US-09-902-481A-3
; Sequence 3, Application US/09902481A
; Publication No. US2003005440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RPT/RMS/RXK
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481A-3

Query Match 100.0%; Score 5879; DB 10; Length 1137;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	FNLDTENAMTFQENARGFGQSVVQGGSRVVVGAPQEIIVAAVNRGSLVQCDYSTGSCBPI 60
Db	1	FNLDTENAMTFQENARGFGQSVVQGGSRVVVGAPQEIIVAAVNRGSLVQCDYSTGSCBPI 60
Qy	61	RLQVPEAVNMSLGLSLAATSPPOLLAGCTVHTCSENTYVKGCLCFGLFGLNLRQOPK 120
Db	61	RLQVPEAVNMSLGLSLAATSPPOLLAGCTVHTCSENTYVKGCLCFGLFGLNLRQOPK 120
Qy	121	PFELRGCPQEDSDIAFLIDSGSIIIPDPRMKELVSTIMEQLKSKTFLPSLMQYSEEF 180

Db 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRRMKELVSTIMEQLKSKTFLSLFMOYSEBF 180
Qy 181 RIHFTKEFQNNPNSLIKPIITQLGRHTATGLRKVVRELFTNTNGARKNAFKILFL 240
Db 181 RIHFTKEFQNNPNSLIKPIITQLGRHTATGLRKVVRELFTNTNGARKNAFKILFL 240
Qy 241 TDEKFGDPLGYEDVPELDRGVIRYVLFQDFAFSEKSRQELNTVASKPRDHVFQAN 300
Db 241 TDEKFGDPLGYEDVPELDRGVIRYVLFQDFAFSEKSRQELNTVASKPRDHVFQAN 300
Qy 301 NFEALKTQNLREKIPAIETGTOTGSSSPSEHMSQEGFSAITSNGLPILSTVGSVDWAG 360
Db 301 NFEALKTQNLREKIPAIETGTOTGSSSPSEHMSQEGFSAITSNGLPILSTVGSVDWAG 360
Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAIILRNRVQSLVLAGPYQHIGLVAMFR 420
Db 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAIILRNRVQSLVLAGPYQHIGLVAMFR 420
Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYTEQTRGGQVSVCP 480
Db 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYTEQTRGGQVSVCP 480
Qy 481 PRGORARWQCDVAVLYGEGQCPWGRFGAALTVDVANGDKLTDVAIGAPGEEDNRGAIVLF 540
Db 481 PRGORARWQCDVAVLYGEGQCPWGRFGAALTVDVANGDKLTDVAIGAPGEEDNRGAIVLF 540
Qy 541 HGTSGSGISPSHSORLAGSKLSPRLQYFGOSLGGQDLTMDGLVDTVGAQGHVLLRSQ 600
Db 541 HGTSGSGISPSHSORLAGSKLSPRLQYFGOSLGGQDLTMDGLVDTVGAQGHVLLRSQ 600
Qy 601 PVLRVKAIMFEFPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLRSGQISVVT 660
Db 601 PVLRVKAIMFEFPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLRSGQISVVT 660
Qy 661 YDLALDSGRPHSRVAFNETKNSRTRQTVLGLTQCTETLKLQLENCEIDPVSPIVLRNF 720
Db 661 YDLALDSGRPHSRVAFNETKNSRTRQTVLGLTQCTETLKLQLENCEIDPVSPIVLRNF 720
Qy 721 SLVGTPLSAFNGLRPVLAEDAQRFLTALFPPEKNCNDNICQDDLSITTFSPMSLDCLVVG 780
Db 721 SLVGTPLSAFNGLRPVLAEDAQRFLTALFPPEKNCNDNICQDDLSITTFSPMSLDCLVVG 780
Qy 781 GPREFNVTVVRNDGDSYRTQVTFPPDLDSYRKVSTLQNSQSRWELACESASSTEV 840
Db 781 GPREFNVTVVRNDGDSYRTQVTFPPDLDSYRKVSTLQNSQSRWELACESASSTEV 840
Qy 841 SGALKSTCSINHPIPEKSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
Db 841 SGALKSTCSINHPIPEKSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
Qy 901 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMQHOYQVSNLQORSPLISLVPLVPV 960
Db 901 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMQHOYQVSNLQORSPLISLVPLVPV 960
Qy 961 RLNQTVINDRPOVTFSENLSSTCHTERLPSSDFLAELRKAPVNCSTAVCORIQCDIP 1020
Db 961 RLNQTVINDRPOVTFSENLSSTCHTERLPSSDFLAELRKAPVNCSTAVCORIQCDIP 1020
Qy 1021 PFGIQEENATLKNLSFDWYIKTSHNLLIVSTABILEFNDVFTLLPQGAFAVRSQET 1080
Db 1021 PFGIQEENATLKNLSFDWYIKTSHNLLIVSTABILEFNDVFTLLPQGAFAVRSQET 1080
Qy 1081 KVEPPEVPNPLPIVGVSSVGGILLALITAAALYKLGFFRQYKQDMMSGGPPGABPQ 1137
Db 1081 KVEPPEVPNPLPIVGVSSVGGILLALITAAALYKLGFFRQYKQDMMSGGPPGABPQ 1137

RESULT 2

US-09-902-481A-4

; Sequence 4, Application US/09902481A

; Publication No. US2003005440A1

; GENERAL INFORMATION:

; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481A-4

Query Match 99.6%; Score 5857; DB 10; Length 1137;

Best Local Similarity 99.5%; Fred. No. 0;

Matches 1131; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTQENARFGQSVVOLQGSRRVVVGAPOEIVAAANQRGSLYCCDYSTGSCBPI 60

Db 1 FNLDTENAMTQENARFGQSVVOLQGSRRVVVGAPOEIVAAANQRGSLYCCDYSTGSCBPI 60

Qy 61 RLQVPVEAVNNSLGLSLAATTSPPLLACGPTVHTQTSENTYVKGCLCFGLFNSNLRQOPK 120

Db 61 RLQVPVEAVNNSLGLSLAATTSPPLLACGPTVHTQTSENTYVKGCLCFGLFNSNLRQOPK 120

Qy 121 PPEALRGCPQEDSDIAFLIDGSGSIIPHDPRRMKELVSTIMEQLKSKTFLSLMOYSEBF 180

Db 121 PPEALRGCPQEDSDIAFLIDGSGSIIPHDPRRMKELVSTIMEQLKSKTFLSLMOYSEBF 180

Qy 181 RIHFTKEFQNNPNSLIKPIITQLGRHTATGLRKVVRELFTNTNGARKNAFKILFL 240

Db 181 RIHFTKEFQNNPNSLIKPIITQLGRHTATGLRKVVRELFTNTNGARKNAFKILFL 240

Qy 241 TDGKFGDPLGYEDVPELDRGVIRYVLFQDFAFSEKSRQELNTVASKPRDHVFQAN 300

Db 241 TDGKFGDPLGYEDVPELDRGVIRYVLFQDFAFSEKSRQELNTVASKPRDHVFQAN 300

Qy 301 NFEALKTQNLREKIPAIETGTOTGSSSPSEHMSQEGFSAITSNGLPILSTVGSVDWAG 360

Db 301 NFEALKTQNLREKIPAIETGTOTGSSSPSEHMSQEGFSAITSNGLPILSTVGSVDWAG 360

Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAIILRNRVQSLVLAGPYQHIGLVAMFR 420

Db 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAIILRNRVQSLVLAGPYQHIGLVAMFR 420

Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYTEQTRGGQVSVCP 480

Db 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYTEQTRGGQVSVCP 480

Qy 481 PRGORARWQCDVAVLYGEGQCPWGRFGAALTVDVANGDKLTDVAIGAPGEEDNRGAIVLF 540

Db 481 PRGORARWQCDVAVLYGEGQCPWGRFGAALTVDVANGDKLTDVAIGAPGEEDNRGAIVLF 540

Qy 541 HGTSGSGISPSHSORLAGSKLSPRLQYFGOSLGGQDLTMDGLVDTVGAQGHVLLRSQ 600

Db 541 HGTSGSGISPSHSORLAGSKLSPRLQYFGOSLGGQDLTMDGLVDTVGAQGHVLLRSQ 600

Qy 601 PVLRVKAIMFEFPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLRSGQISVVT 660

Db 601 PVLRVKAIMFEFPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLRSGQISVVT 660

Qy 661 YDLALDSGRPHSRVAFNETKNSRTRQTVLGLTQCTETLKLQLENCEIDPVSPIVLRNF 720

Db 661 YDLALDSGRPHSRVAFNETKNSRTRQTVLGLTQCTETLKLQLENCEIDPVSPIVLRNF 720

Qy 721 SLVGTPLSAFNGLRPVLAEDAQRFLTALFPPEKNCNDNICQDDLSITTFSPMSLDCLVVG 780

Db 721 SVUGTSLAFGNLRPVLAEDAQRLEFALPFPFKKNGNDNICQDDLSITFSFMSDCLVVG 780
Qy 781 GPREFNVTIVRNDGEDSYRTQVTFPFPDLDSYRKVSTIQNORSORSWRLACESASSTEV 840
Db 781 GPREFNVTIVRNDGEDSYRTQVTFPFPDLDSYRKVSTIQNORSORSWRLACESASSTEV 840
Qy 841 SGALKSTSCSINHPIPPENSEVTFTNITFDVDSKASLGKLLKXNVTSNNMPTNKTEP 900
Db 841 SGALKSTSCSINHPIPPENSEVTFTNITFDVDSKASLGKLLKXNVTSNNMPTNKTEP 900
Qy 901 QLELPVKYAVYVMTSHGVSTKYLNFTASENTRVQHQYQVSNLQORSLSPLSLVFLVPV 960
Db 901 QLELPVKYAVYVMTSHGVSTKYLNFTASENTRVQHQYQVSNLQORSLSPLSLVFLVPV 960
Qy 961 RLNQTVIWDREPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVYVNCIAVCQRIQCDDIP 1020
Db 961 RLNQTVIWDREPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVYVNCIAVCQRIQCDDIP 1020
Qy 1021 FFGIOBEFNATLKGKNSLSPDNYIKTSHNLLIYSTAEIILFNDSDVFTLLPQOGAFVRSQDET 1080
Db 1021 FFGIOBEFNATLKGKNSLSPDNYIKTSHNLLIYSTAEIILFNDSDVFTLLPQOGAFVRSQDET 1080
Qy 1081 KVEPPEVNPPLPLIVGSSVGGLLALITAAALYKLGFPKQYKDMWSEGGPPGABPQ 1137
Db 1081 KVEPPEVNPPLPLIVGSSVGGLLALITAAALYKLGFPKQYKDMWSEGGPPGABPQ 1137

RESULT 3

US-09-350-259-3
; Sequence 3, Application US/09350259
; Patent No. US20020062008A1
; GENERAL INFORMATION:

; APPLICANT: Gallatin, Michael W.
; TITLE OF INVENTION: No. US20020062008A1 Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350,259
; EARLIER FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: 09/193,043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 3
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-259-3

Query Match 99.3%; Score 5839; DB 9; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

Qy 1 ENLDTENATVFOENARGFGQSVVUQGSRVVVGAPQEIIVANQKSLVQCDYSTGSCRP 60
Db 17 FNLDTENATVFOENARGFGQSVVUQGSRVVVGAPQEIIVANQKSLVQCDYSTGSCRP 76
Qy 61 RLQVPVEAVNMSLGLSLAATSPQLACGPTVHOTCSENTYVKGCLFLGSLNLRQOPQK 120
Db 77 RLQVPVEAVNMSLGLSLAATSPQLACGPTVHOTCSENTYVKGCLFLGSLNLRQOPQK 136
Qy 121 FPEALRGCPQSDIAFLDGSGLIIPDPRFMKELVSTIMEOLKSKTLPGLMOYSEF 180
Db 137 FPEALRGCPQSDIAFLDGSGLIIPDPRFMKELVSTIMEOLKSKTLPGLMOYSEF 196

Qy 181 RIHETFKEPQNNPRLSKIPITOLLGRTHATGLRKVWVRELNFNTNGARNAKLELL 240
Db 197 RIHETFKEPQNNPRLSKIPITOLLGRTHATGLRKVWVRELNFNTNGARNAKLELL 256
Qy 241 TDGEKFGDPLGYDVIDPELQREGVIRVVLGPDGAFRSEKSRQBLNTVASKPFRDHVQAN 300
Db 257 TDGEKFGDPLGYDVIDPELQREGVIRVVLGPDGAFRSEKSRQBLNTVASKPFRDHVQAN 316
Qy 301 NFEALTKVQMLREKIIFAIETGOTGSSSPHEHKSQSPSAITNSNGELLSTVGSYDWAG 360
Db 317 NFEALTKVQMLREKIIFAIETGOTGSSSPHEHKSQSPSAITNSNGELLSTVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFTNMTTRVDSMDNDAYLGVAALIIILNRVQSLVGLAPRYOHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFTNMTTRVDSMDNDAYLGVAALIIILNRVQSLVGLAPRYOHIGLVAMFR 436
Qy 421 QNTGWSNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYETQTRGGQSVCP 480
Db 437 QNTGWSNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYETQTRGGQSVCP 496
Qy 481 PRGORARWQCDVAVLYGQGPWGRFGAALTIVLGVNDGKLTVAIGAPGEEDNKGAVYLF 540
Db 497 PRGORARWQCDVAVLYGQGPWGRFGAALTIVLGVNDGKLTVAIGAPGEEDNKGAVYLF 556
Qy 541 HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGQDLTMDGLVDLTVGAQSHVLLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGQDLTMDGLVDLTVGAQSHVLLRSQ 616
Qy 601 PVLRAVKALMBENPREVARNVFECDQVVKGEAGEVRVCLHVOKSTRDLRSGQSVVT 660
Db 617 PVLRAVKALMBENPREVARNVFECDQVVKGEAGEVRVCLHVOKSTRDLRSGQSVVT 676
Qy 661 YDLALDSGRPHSRVAFNETKNSRTRQVGLTQTCETLKLQLPNCIEDPVSPIVRLNF 720
Db 677 YDLALDSGRPHSRVAFNETKNSRTRQVGLTQTCETLKLQLPNCIEDPVSPIVRLNF 736
Qy 721 SLVGTPLSAFQNLRPVLAEDAQRLEFALPFPFKKNGNDNICQDDLSITFSFMSDCLVVG 780
Db 737 SLVGTPLSAFQNLRPVLAEDAQRLEFALPFPFKKNGNDNICQDDLSITFSFMSDCLVVG 796
Qy 781 GPREFNVTIVRNDGEDSYRTQVTFPFPDLDSYRKVSTIQNORSORSWRLACESASSTEV 840
Db 797 GPREFNVTIVRNDGEDSYRTQVTFPFPDLDSYRKVSTIQNORSORSWRLACESASSTEV 856
Qy 841 SGALKSTSCSINHPIPPENSEVTFTNITFDVDSKASLGKLLKXNVTSNNMPTNKTEP 900
Db 857 SGALKSTSCSINHPIPPENSEVTFTNITFDVDSKASLGKLLKXNVTSNNMPTNKTEP 916
Qy 901 QLELPVKYAVYVMTSHGVSTKYLNFTASENTRVQHQYQVSNLQORSLSPLSLVFLVPV 960
Db 917 QLELPVKYAVYVMTSHGVSTKYLNFTASENTRVQHQYQVSNLQORSLSPLSLVFLVPV 976
Qy 961 RLNQTVIWDREPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVYVNCIAVCQRIQCDDIP 1020
Db 977 RLNQTVIWDREPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVYVNCIAVCQRIQCDDIP 1036
Qy 1021 FFGIOBEFNATLKGKNSLSPDNYIKTSHNLLIYSTAEIILFNDSDVFTLLPQOGAFVRSQDET 1080
Db 1037 FFGIOBEFNATLKGKNSLSPDNYIKTSHNLLIYSTAEIILFNDSDVFTLLPQOGAFVRSQDET 1096
Qy 1081 KVEPPEVNPPLPLIVGSSVGGLLALITAAALYKLGFPKQYKDMWSEGGPPGABPQ 1137
Db 1097 KVEPPEVNPPLPLIVGSSVGGLLALITAAALYKLGFPKQYKDMWSEGGPPGABPQ 1153

RESULT 4

US-09-902-481A-1
; Sequence 1, Application US/09902481A
; Publication No. US20030054440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimooka, Motomu
; APPLICANT: Shifman, Julia

APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/FT/IRMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: mat peptide
; LOCATION: (17)...()
; OTHER INFORMATION:
US-09-902-481A-1

Query Match 99.3%; Score 5839; DB 10; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARGFQGVVQLQGSRRVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFOENARGFQGVVQLQGSRRVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI 76
QY 61 RLQVPVEAVNMVSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGCLPFGSLNLRQPOK 120
DB 77 RLQVPVEAVNMVSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGCLPFGSLNLRQPOK 136
QY 121 FPEALRGCPQSDSDIAFLIDGSGSIIPHPRMKELVSTIMEOLKSKTLPFSLMOYSEEF 180
DB 137 FPEALRGCPQSDSDIAFLIDGSGSIIPHPRMKELVSTIMEOLKSKTLPFSLMOYSEEF 196
QY 181 RIHFTPEFQNNPRSLIKPITQLGRTHATGLRKVVRELFNITNGARKNAKILFLL 240
DB 197 RIHFTPEFQNNPRSLIKPITQLGRTHATGLRKVVRELFNITNGARKNAKILFLL 256
QY 241 TDSEKPGDPLGYEDVPELDREGVIRYVLGFGDAFSEKSRQELNVTASKPRDHVFQAN 300
DB 257 TDSEKPGDPLGYEDVPELDREGVIRYVLGFGDAFSEKSRQELNVTASKPRDHVFQAN 316
QY 301 NPEALTVONLRKIFATECTQTGSSSFHEHMSQEGFSAITNSGPLLSTVGSYDNAG 360
DB 317 NPEALTVONLRKIFATECTQTGSSSFHEHMSQEGFSAITNSGPLLSTVGSYDNAG 376
QY 361 GVFLYTSKEKSTPINKTRVDSMDNDAYLGAAAILLRNVQSLVLCAPRYQHIGLVAMFR 420
DB 377 GVFLYTSKEKSTPINKTRVDSMDNDAYLGAAAILLRNVQSLVLCAPRYQHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYETRGQGVSVCP 480
DB 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYETRGQGVSVCP 496
QY 481 PRQARWQCDVLYEQCPWRCFGAALTLDVNGDKLTDVAIGAPGEENRGAVILF 540
DB 497 PRQARWQCDVLYEQCPWRCFGAALTLDVNGDKLTDVAIGAPGEENRGAVILF 556
QY 541 HGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLGGQDLTMDGLVLTGCAQGHVLLRSQ 600
DB 557 HGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLGGQDLTMDGLVLTGCAQGHVLLRSQ 616
QY 601 PVLRVKAINFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
DB 617 PVLRVKAINFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676
QY 661 YDLALDSGRPHSAVENETKNSTRQTQVLGTLQTCETLKLQPNCTEDPVSPIVLRNPF 720
DB 677 YDLALDSGRPHSAVENETKNSTRQTQVLGTLQTCETLKLQPNCTEDPVSPIVLRNPF 736
QY 721 SLVGTPLSAFGNLRPVLAEDAQLFTALFPPEKNCNDNICODDLSITFSPMSLDCLVVG 780

DB 737 SLVGTPLSAFGNLRPVLAEDAQLFTALFPPEKNCNDNICODDLSITFSPMSLDCLVVG 796
QY 781 GPRBFNVTVVRNDEGDSYRTQVTFFPPLDLISYRKVSTLQNRQSRWRLACESASSTEV 840
DB 797 GPRBFNVTVVRNDEGDSYRTQVTFFPPLDLISYRKVSTLQNRQSRWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
DB 857 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 916
QY 901 QLELPVKYAVVMVTSBGVSTKYLNFTASENTRSMOHQYQVSNLQGSRLPISLVFLVPV 960
DB 917 QLELPVKYAVVMVTSBGVSTKYLNFTASENTRSMOHQYQVSNLQGSRLPISLVFLVPV 976
QY 961 RLNQTIVWDRPQVTFSENLSSSTCHTKERLPSHSDPLAELRKAPVVCNCSIAVCORIQCDIP 1020
DB 977 RLNQTIVWDRPQVTFSENLSSSTCHTKERLPSHSDPLAELRKAPVVCNCSIAVCORIQCDIP 1036
QY 1021 PFGIOEEFNATLKGKLSFDMYIKTSHNLLIVSTABILFNDVSFTLLPGQAGAFVRSQTEF 1080
DB 1037 PFGIOEEFNATLKGKLSFDMYIKTSHNLLIVSTABILFNDVSFTLLPGQAGAFVRSQTEF 1096
QY 1081 KYEPFEPVNPPLTVGSSVGGLLILALITAAALYKLGPKROYKDMMSGGPPGAEPPQ 1137
DB 1097 KYEPFEPVNPPLTVGSSVGGLLILALITAAALYKLGPKROYKDMMSGGPPGAEPPQ 1153

RESULT 5
US-09-891-943-3
; Sequence 3, Application US/09891943
; Publication No. US2003007278A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US2003007278a1 Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/891,943
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-891-943-3

Query Match 99.3%; Score 5839; DB 10; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARGFQGVVQLQGSRRVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFOENARGFQGVVQLQGSRRVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI 76
QY 61 RLQVPVEAVNMVSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGCLPFGSLNLRQPOK 120
DB 77 RLQVPVEAVNMVSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGCLPFGSLNLRQPOK 136
QY 121 FPEALRGCPQSDSDIAFLIDGSGSIIPHPRMKELVSTIMEOLKSKTLPFSLMOYSEEF 180
DB 137 FPEALRGCPQSDSDIAFLIDGSGSIIPHPRMKELVSTIMEOLKSKTLPFSLMOYSEEF 196
QY 181 RIHFTPEFQNNPRSLIKPITQLGRTHATGLRKVVRELFNITNGARKNAKILFLL 240
DB 197 RIHFTPEFQNNPRSLIKPITQLGRTHATGLRKVVRELFNITNGARKNAKILFLL 256


```
QY 841 SGALKSTSCSINHPIEPENSEVTENITFDVDSKASLGNKLLKANVTSENNMPTNKTEP 900
DB 857 SGALKSTSCSINHPIEPENSEVTENITFDVDSKASLGNKLLKANVTSENNMPTNKTEP 916
QY 901 QLELPVKYAVYVMVTSHGVSSTKYLNFTASENTSRVMQHOYQVSNLQORSLSPLSLVFLVPV 960
DB 917 QLELPVKYAVYVMVTSHGVSSTKYLNFTASENTSRVMQHOYQVSNLQORSLSPLSLVFLVPV 976
QY 961 RLNOTVIWDRPQVTFSENISSCTCHTKERLPSSHDFLAELRKAPVWNCSTAVCQRIQCDDIP 1020
DB 977 RLNOTVIWDRPQVTFSENISSCTCHTKERLPSSHDFLAELRKAPVWNCSTAVCQRIQCDDIP 1036
QY 1021 FFGIOEFPNATLKGNSLFDWYIKTSHNHLIYSTAEILFNDVSFTLLPQCGAFVRSQTEP 1080
DB 1037 FFGIOEFPNATLKGNSLFDWYIKTSHNHLIYSTAEILFNDVSFTLLPQCGAFVRSQTEP 1096
QY 1081 KVEPPEVPNPLPLIVGSSVGGLLLLALITAAALYKLGFFKQYKDMMSGGPPGAEPPQ 1137
DB 1097 KVEPPEVPNPLPLIVGSSVGGLLLLALITAAALYKLGFFKQYKDMMSGGPPGAEPPQ 1153

RESULT 7
US-10-207-655-176
; Sequence 176, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Hayden-Ledbetter, Martha S.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069.401C1
; CURRENT APPLICATION NUMBER: US/10/207,655
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 176
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-207-655-176

Query Match 99.3%; Score 5839; DB 14; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 FNLDTENATTFQENARSGQSVVQLGSRVWVGAPQEIIVAAANQORSLYQCDYSTGSCERI 60
DB 17 FNLDTENATTFQENARSGQSVVQLGSRVWVGAPQEIIVAAANQORSLYQCDYSTGSCERI 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHOTCSENTYVYKGLCFGLFGSNLRQPOK 120
DB 77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHOTCSENTYVYKGLCFGLFGSNLRQPOK 136
QY 121 FPEALGCGQEDSDIAFLIDGSGIIPHPFRMKELVSTIMEQLKSKTFLSLMOYSEEP 180
DB 137 FPEALGCGQEDSDIAFLIDGSGIIPHPFRMKELVSTIMEQLKSKTFLSLMOYSEEP 196
QY 181 RLHFTKEQNNPNRSLKIPITOLLGRTHATGIRKVVRELFTNNGARKNAKILPL 240
DB 197 RLHFTKEQNNPNRSLKIPITOLLGRTHATGIRKVVRELFTNNGARKNAKILPL 256
QY 241 TGEKFGDPLGYEDVYIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPRDHYFQAN 300
DB 257 TGEKFGDPLGYEDVYIPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPRDHYFQAN 316
QY 301 NPEALKTQNLREKIFALBGTGTGSSSEFHEMSQEGFSAATSNGLPILSTVGSVDWAG 360
DB 317 NPEALKTQNLREKIFALBGTGTGSSSEFHEMSQEGFSAATSNGLPILSTVGSVDWAG 376
QY 361 GVFLYTSKSKSTFINMTRVDSMDNDAYLGAAAILLRNRVQSLVGLGAPRYQHIGLVAMER 420
DB 377 GVFLYTSKSKSTFINMTRVDSMDNDAYLGAAAILLRNRVQSLVGLGAPRYQHIGLVAMER 436
QY 421 QNTGWMESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTGGQVSVCP 480
```

```
DB 437 QNTGWMESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTGGQVSVCP 496
QY 481 PRGQARWQCDVLYGEOQPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNRGAVYLF 540
DB 497 PRGQARWQCDVLYGEOQPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNRGAVYLF 556
QY 541 HGTSGSGISPSHSORLIAGSKLSPLOYGQSLSGQDITMDGLVDLTVGAQGHVLLRSQ 600
DB 557 HGTSGSGISPSHSORLIAGSKLSPLOYGQSLSGQDITMDGLVDLTVGAQGHVLLRSQ 616
QY 601 PVLAVKAIMFNPFRVARNVFECDQVYVKGKAGEVRVCLHVQKSTRDELREGQISVYT 660
DB 617 PVLAVKAIMFNPFRVARNVFECDQVYVKGKAGEVRVCLHVQKSTRDELREGQISVYT 676
QY 661 YDLALDSRPHSRVAFNETKNSRTOYVGLTGTCTEKLQLPNCIEDPSPVILRLNP 720
DB 677 YDLALDSRPHSRVAFNETKNSRTOYVGLTGTCTEKLQLPNCIEDPSPVILRLNP 736
QY 721 SLVGTPLSAFCNLAPVLAEDAQRLLFTALFPPEKKNCGNDNICODDLSTIFFSMSLDCLVVG 780
DB 737 SLVGTPLSAFCNLAPVLAEDAQRLLFTALFPPEKKNCGNDNICODDLSTIFFSMSLDCLVVG 796
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
DB 797 GPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIEPENSEVTENITFDVDSKASLGNKLLKANVTSENNMPTNKTEP 900
DB 857 SGALKSTSCSINHPIEPENSEVTENITFDVDSKASLGNKLLKANVTSENNMPTNKTEP 916
QY 901 QLELPVKYAVYVMVTSHGVSSTKYLNFTASENTSRVMQHOYQVSNLQORSLSPLSLVFLVPV 960
DB 917 QLELPVKYAVYVMVTSHGVSSTKYLNFTASENTSRVMQHOYQVSNLQORSLSPLSLVFLVPV 976
QY 961 RLNOTVIWDRPQVTFSENISSCTCHTKERLPSSHDFLAELRKAPVWNCSTAVCQRIQCDDIP 1020
DB 977 RLNOTVIWDRPQVTFSENISSCTCHTKERLPSSHDFLAELRKAPVWNCSTAVCQRIQCDDIP 1036
QY 1021 FFGIOEFPNATLKGNSLFDWYIKTSHNHLIYSTAEILFNDVSFTLLPQCGAFVRSQTEP 1080
DB 1037 FFGIOEFPNATLKGNSLFDWYIKTSHNHLIYSTAEILFNDVSFTLLPQCGAFVRSQTEP 1096
QY 1081 KVEPPEVPNPLPLIVGSSVGGLLLLALITAAALYKLGFFKQYKDMMSGGPPGAEPPQ 1137
DB 1097 KVEPPEVPNPLPLIVGSSVGGLLLLALITAAALYKLGFFKQYKDMMSGGPPGAEPPQ 1153

RESULT 8
US-09-902-481A-5
; Sequence 5, Application US/09902481A
; Publication No. US2003005440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RPT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481A-5
```

Query Match 99.3%; Score 5836; DB 10; Length 1137;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 1 FNLDTENAMTQENARFGQSVVQLQGSRVVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI 60
DB 1 FNLDTENAMTQENARFGQSVVQLQGSRVVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI 60

QY 61 RLOVPVEAVNMSLGLSLAATTPPQLACGPTVHOTCSENTYVKGCLFGLSGNLQOQOK 120
DB 61 RLOVPVEAVNMSLGLSLAATTPPQLACGPTVHOTCSENTYVKGCLFGLSGNLQOQOK 120

QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKELVSTIMEQLKSKTFLSLMOYSSEF 180
DB 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKELVSTIMEQLKSKTFLSLMOYSSEF 180

QY 181 RIHFTFKFQNNPNRSLIKPTOLLGRTHRTATGLRKVVRELENTNGARKNAFKILFL 240
DB 181 RIHFTFKFQNNPNRSLIKPTOLLGRTHRTATGLRKVVRELENTNGARKNAFKILFL 240

QY 241 TDGEKFGDPLGYEDVIPEDREGVIRYVVGDAFSEKSRQELNTVASKPPDRHVFOAN 300
DB 241 TDGEKFGDPLGYEDVIPEDREGVIRYVVGDAFSEKSRQELNTVASKPPDRHVFOAN 300

QY 301 NFEALKTQVQNLREKIIPHDPRMKELVSTIMEQLKSKTFLSLMOYSSEF 360
DB 301 NFEALKTQVQNLREKIIPHDPRMKELVSTIMEQLKSKTFLSLMOYSSEF 360

QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAILLRNVQSLVGLGAPRYOHIGLVAMFR 420
DB 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAILLRNVQSLVGLGAPRYOHIGLVAMFR 420

QY 421 QNTGWESENAVKTQIGAYFGASLCSDVDVDSNGSTDVLIGAPHYYQOTGGQSVCP 480
DB 421 QNTGWESENAVKTQIGAYFGASLCSDVDVDSNGSTDVLIGAPHYYQOTGGQSVCP 480

QY 481 PRGORARWQCDAVLYGEOQPGWRFGAALTVDVNGDKLTDVAIGAPGEEDNRCVYLF 540
DB 481 PRGORARWQCDAVLYGEOQPGWRFGAALTVDVNGDKLTDVAIGAPGEEDNRCVYLF 540

QY 541 HGTSGSGISPSHSQRIAGSKLSPLOYFGQSLGGQDLTMDGLVDTVGAQGHVLLRSQ 600
DB 541 HGTSGSGISPSHSQRIAGSKLSPLOYFGQSLGGQDLTMDGLVDTVGAQGHVLLRSQ 600

QY 601 PVLRVKAIIMEFNPREVARNVFECNDQVXGKAGEVRVCLHVQKSTRDLRGOIQSVVT 660
DB 601 PVLRVKAIIMEFNPREVARNVFECNDQVXGKAGEVRVCLHVQKSTRDLRGOIQSVVT 660

QY 661 YDLALDSGRPHSRVAFNETKSTRTOVLGTQTCETLKLQPCIEDPVSPVLRNF 720
DB 661 YDLALDSGRPHSRVAFNETKSTRTOVLGTQTCETLKLQPCIEDPVSPVLRNF 720

QY 721 SLVGTPLSAPGNLPRVLAEDAQLEFALPFPFKNCGNDNICQDDLSITPFSMSLCLVVG 780
DB 721 SLVGTPLSAPGNLPRVLAEDAQLEFALPFPFKNCGNDNICQDDLSITPFSMSLCLVVG 780

QY 781 GREFNVTYVRNDGEDSVRTQVTPFPFLDLSVRKUSTLQNRORSRWELACESASSTEV 840
DB 781 GREFNVTYVRNDGEDSVRTQVTPFPFLDLSVRKUSTLQNRORSRWELACESASSTEV 840

QY 841 SGALKSTSCSINHPIFPENSEVFNITPDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
DB 841 SGALKSTSCSINHPIFPENSEVFNITPDVDSKASLGNKLLKANVTSENMPRTNKTEF 900

QY 901 QLELPVKYAVMYVTSKGVSTKYNLTASENTSRVWQHGYQVSNIGORSLSPLSLVFLVPV 960
DB 901 QLELPVKYAVMYVTSKGVSTKYNLTASENTSRVWQHGYQVSNIGORSLSPLSLVFLVPV 960

QY 961 RLNOTVINDRPOVTFSENLSSTCHTKERLPSSDFLAELRKPVPVNCSTAVCORQCDDIP 1020
DB 961 RLNOTVINDRPOVTFSENLSSTCHTKERLPSSDFLAELRKPVPVNCSTAVCORQCDDIP 1020

QY 1021 PFGIOEEFNATLKGNSLFDWYIKTSINHLIIVSTAEILFNDVSFTLLPQCGAFVRSQTEF 1080

DB 1021 PFGIOEEFNATLKGNSLFDWYIKTSINHLIIVSTAEILFNDVSFTLLPQCGAFVRSQTEF 1080
QY 1081 KVEPFEVNPPLPLVGGSVGGLLALLTAALYKLGPFKROYKMMSEGGPPGABPQ 1137
DB 1081 KVEPFEVNPPLPLVGGSVGGLLALLTAALYKLGPFKROYKMMSEGGPPGABPQ 1137

RESULT 9
US-09-902-481A-6
; Sequence 6, Application US/09902481A
; Publication No. US2003005440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RPT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 6
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481A-6

Query Match 99.2%; Score 5832; DB 10; Length 1137;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 1122; Conservative 9; Mismatches 6; Indels 0; Gaps 0;

QY 1 FNLDTENAMTQENARFGQSVVQLQGSRVVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI 60
DB 1 FNLDTENAMTQENARFGQSVVQLQGSRVVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI 60

QY 61 RLOVPVEAVNMSLGLSLAATTPPQLACGPTVHOTCSENTYVKGCLFGLSGNLQOQOK 120
DB 61 RLOVPVEAVNMSLGLSLAATTPPQLACGPTVHOTCSENTYVKGCLFGLSGNLQOQOK 120

QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKELVSTIMEQLKSKTFLSLMOYSSEF 180
DB 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKELVSTIMEQLKSKTFLSLMOYSSEF 180

QY 181 RIHFTFKFQNNPNRSLIKPTOLLGRTHRTATGLRKVVRELENTNGARKNAFKILFL 240
DB 181 RIHFTFKFQNNPNRSLIKPTOLLGRTHRTATGLRKVVRELENTNGARKNAFKILFL 240

QY 241 TDGEKFGDPLGYEDVIPEDREGVIRYVVGDAFSEKSRQELNTVASKPPDRHVFOAN 300
DB 241 TDGEKFGDPLGYEDVIPEDREGVIRYVVGDAFSEKSRQELNTVASKPPDRHVFOAN 300

QY 301 NFEALKTQVQNLREKIIPHDPRMKELVSTIMEQLKSKTFLSLMOYSSEF 360
DB 301 NFEALKTQVQNLREKIIPHDPRMKELVSTIMEQLKSKTFLSLMOYSSEF 360

QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAILLRNVQSLVGLGAPRYOHIGLVAMFR 420
DB 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAILLRNVQSLVGLGAPRYOHIGLVAMFR 420

QY 421 QNTGWESENAVKTQIGAYFGASLCSDVDVDSNGSTDVLIGAPHYYQOTGGQSVCP 480
DB 421 QNTGWESENAVKTQIGAYFGASLCSDVDVDSNGSTDVLIGAPHYYQOTGGQSVCP 480

QY 481 PRGORARWQCDAVLYGEOQPGWRFGAALTVDVNGDKLTDVAIGAPGEEDNRCVYLF 540
DB 481 PRGORARWQCDAVLYGEOQPGWRFGAALTVDVNGDKLTDVAIGAPGEEDNRCVYLF 540

```
QY 541 HGTSGSISPSHSORISAGSKLSPRLQYFGOSLSGGDLTMDGLVDLTGAGQGHVLLRSQ 600
DB 541 HGTSGSISPSHSORISAGSKLSPRLQYFGOSLSGGDLTMDGLVDLTGAGQGHVLLRSQ 600
QY 601 PVLRVKAIMENFPREVARNVFECNDQVVKGKEAGEVRLVQKSTRDLRREGQIQSVVT 660
DB 601 PVLRVKAIMENFPREVARNVFECNDQVVKGKEAGEVRLVQKSTRDLRREGQIQSVVT 660
QY 661 YDLALDSGRPHSRAPVNETKSTRQOTQVLGLTQTCETKLQLPNCIEDPVSPIVRLNPF 720
DB 661 YDLALDSGRPHSRAPVNETKSTRQOTQVLGLTQTCETKLQLPNCIEDPVSPIVRLNPF 720
QY 721 SLVGTPLSAFENLRPVLAEDAQRLLFTALPPFEKNCNDNI CODDLSITFSFMSLDCLVVG 780
DB 721 SLVGTPLSAFENLRPVLAEDAQRLLFTALPPFEKNCNDNI CODDLSITFSFMSLDCLVVG 780
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNSORSWRLACESASSTEV 840
DB 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNSORSWRLACESASSTEV 840
QY 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGKLLKANVTSENMPRTNKTEP 900
DB 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGKLLKANVTSENMPRTNKTEP 900
QY 901 QLELPVKYAVYVMVTSIGVSTKYNFTASENTSRVMQHYQVSNLQORSPLISLVLVPV 960
DB 901 QLELPVKYAVYVMVTSIGVSTKYNFTASENTSRVMQHYQVSNLQORSPLISLVLVPV 960
QY 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVKCSIAVCORIQCDIP 1020
DB 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVKCSIAVCORIQCDIP 1020
QY 1021 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIIVSTABILFNDVSFTLLPGQGFVRSQTET 1080
DB 1021 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIIVSTABILFNDVSFTLLPGQGFVRSQTET 1080
QY 1081 KVEPFEVNPFLPIVGVSSVGLLLALITAALYKLGFFKQYKDMMSBGGPPGABEQ 1137
DB 1081 KVEPFEVNPFLPIVGVSSVGLLLALITAALYKLGFFKQYKDMMSBGGPPGABEQ 1137

RESULT 10
US-09-945-265-4
; Sequence 4, Application US/09945265
; Patent No. US20020123614A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy A.
; APPLICANT: Shimooka, Motomu
; APPLICANT: Lu, Chafen
; TITLE OF INVENTION: MODIFIED POLYPEPTIDES STABILIZED IN A
; FILE REFERENCE: DESIRED CONFORMATION AND METHODS FOR PRODUCING SAME
; CURRENT APPLICATION NUMBER: CRN-002CP
; CURRENT FILING DATE: 2001-08-31
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1152
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-945-265-4

Query Match 99.1%; Score 5823.5; DB 9; Length 1152;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 1123; Conservative 8; Mismatches 5; Indels 1; Gaps 1;

QY 1 FNLDTENAMTQENARFGQSVVOLQGSRRVVGAPQBIIVAAVNRGSLYQCDYSTGSCPT 60
DB 17 FNLDTENAMTQENARFGQSVVOLQGSRRVVGAPQBIIVAAVNRGSLYQCDYSTGSCPT 76
QY 61 RLQVFEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSENTYVKGCLFLGSLNRQOPQK 120
```

RESULT 11

Db 138 LPVSRQCEPQEQDIFVLIDGSGISIRNFATMNFVRAVISQFQEPSTQFSLMFQSNKF 197
Qy 181 RIHFTKEFQNNPNSLKIPIQLGLRTHATGLKRVVRELNFNTNGARKNAKILFLL 240
Db 198 QTHFTFEBERTSNPLSLASVHQLQGYTATAIQNVVHUFHASYGARRDAIKLIVI 257
Qy 241 TQCEKFGDPLGYEDVPELDRGVIVYLGFGDAFRSEKSRQELNVTASKPPDRHVFOAN 300
Db 258 TQCKGSGDSLVDKVPMDAAGIIRYALGVGLAFQNRNSWELNDIASKPSQEHFKVE 317
Qy 301 NFALKTQVQNLREKIFAIEGTOTGSSSPHEHMSQEGFSAATNSGPLLSTVGSVDWAG 360
Db 318 DFDALKDIONLKEKIFAIEGTETISSSSFELEMAQEGFSAVTPDGPVLGAVGFTWSG 377
Qy 361 GVFLYTSKEKSTFTINMTVDSDMDNDAYLGAAAIILNRNVQSLVGLAPRYQHIGLVAMPR 420
Db 378 GAFLYPPNNSPTFINMSQENVMDRDSYLGISTELALMKGVQSLVGLCAPRYQHIGKAVIFI 437
Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQVSVCP 480
Db 438 QVSRQWRKMAKAVIGTQIGSYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQVSVCP 497
Qy 481 PRQBARQCCDANLYXGEOQWGPFGAALTVDGVNGDKLTDVAIGAPGEDNRCAVYLP 540
Db 498 PRGWR-RWMCDAVLYGEOQWGPFGAALTVDGVNGDKLTDVWVIGAPGEENRGAVYLP 556
Qy 541 HGTSGSGISPSHSORLAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 557 HGVLPSPISPSHSORLAGSKLSSRLQYFGQALSGQDLTQDGLVDLAVGARGOVLLRTR 616
Qy 601 PVLRYKALMEFNPREVARNFECDQVWKGAESEVRYCLAVOKSTRDLREGOQSVVT 660
Db 617 PVLWYGVSMQFIPAEIPRSAFECEQVWSEQLVQSNICLYIDKRSKLLGSRDLQSSVT 676
Qy 661 YDLALDSGRPHSRVAFNETKSTRQTOVLGLTQTCETKLQLPNCIEDPVSPIVLRNF 720
Db 677 LDALAPGRLSPRALFQETKRSLSRVRLGKAHCENFNLLFSCVEDSVIPIILRNF 736

RESULT 13

US-09-891-943-4

; Sequence 4, Application US/09891943

; Publication No. US20030077278A1

; GENERAL INFORMATION:

; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20030077278A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/891,943
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1163
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-891-943-4

Query Match 58.2%; Score 3423; DB 10; Length 1163;

Best Local Similarity 60.2%; Pred. No. 2.8e-307;

Matches 680; Conservative 149; Mismatches 294; Indels 6; Gaps 4;

Qy 1 FNLDTENAMTFOENARGFGQSVVOLQGSRRVVVVGAPQEIIVAAANQSGSLYQCDYSTGSCBPI 60
Db 20 FNLDTEELTAFRVDSAGFGSDVVQYANSVVVGAPQKIIANQIGGLYQCGYSTGACBPI 79
Qy 61 RLQVPEAVNNSLGLSLAATSPOLLACGTPTVQTCSENTYVVKGLCTLFGSNLRQOQPK 120
Db 80 GLQVPEAVNNSLGLSLAATSPOLLACGTPTVQTCSENTYVVKGLCTLFGSNLRQOQPK 137
Qy 121 FPEALRGCPQSDIDIAFLIDGSGSIIIPHDPRMKELVSTIMEQLKSKTLFSLMYSBEF 180
Db 138 LPVSRQCEPQEQDIFVLIDGSGISIRNFATMNFVRAVISQFQEPSTQFSLMFQSNKF 197
Qy 181 RIHFTKEFQNNPNSLKIPIQLGLRTHATGLKRVVRELNFNTNGARKNAKILFLL 240
Db 198 QTHFTFEBERTSNPLSLASVHQLQGYTATAIQNVVHUFHASYGARRDAIKLIVI 257
Qy 241 TQCEKFGDPLGYEDVPELDRGVIVYLGFGDAFRSEKSRQELNVTASKPPDRHVFOAN 300
Db 258 TQCKGSGDSLVDKVPMDAAGIIRYALGVGLAFQNRNSWELNDIASKPSQEHFKVE 317
Qy 301 NFALKTQVQNLREKIFAIEGTOTGSSSPHEHMSQEGFSAATNSGPLLSTVGSVDWAG 360
Db 318 DFDALKDIONLKEKIFAIEGTETISSSSFELEMAQEGFSAVTPDGPVLGAVGFTWSG 377
Qy 361 GVFLYTSKEKSTFTINMTVDSDMDNDAYLGAAAIILNRNVQSLVGLAPRYQHIGLVAMPR 420
Db 378 GAFLYPPNNSPTFINMSQENVMDRDSYLGISTELALMKGVQSLVGLCAPRYQHIGKAVIFI 437
Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQVSVCP 480
Db 438 QVSRQWRKMAKAVIGTQIGSYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQVSVCP 497
Qy 481 PRQBARQCCDANLYXGEOQWGPFGAALTVDGVNGDKLTDVAIGAPGEDNRCAVYLP 540
Db 498 PRGWR-RWMCDAVLYGEOQWGPFGAALTVDGVNGDKLTDVWVIGAPGEENRGAVYLP 556
Qy 541 HGTSGSGISPSHSORLAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 557 HGVLPSPISPSHSORLAGSKLSSRLQYFGQALSGQDLTQDGLVDLAVGARGOVLLRTR 616
Qy 601 PVLRYKALMEFNPREVARNFECDQVWKGAESEVRYCLAVOKSTRDLREGOQSVVT 660
Db 617 PVLWYGVSMQFIPAEIPRSAFECEQVWSEQLVQSNICLYIDKRSKLLGSRDLQSSVT 676
Qy 661 YDLALDSGRPHSRVAFNETKSTRQTOVLGLTQTCETKLQLPNCIEDPVSPIVLRNF 720
Db 677 LDALAPGRLSPRALFQETKRSLSRVRLGKAHCENFNLLFSCVEDSVIPIILRNF 736

novel gene within the common cystinosis-causing deletion.";
Genome Res. 10:165-173(2000).
[4]
RX MUTAGENESIS OF ASP-109 AND PHE-316.
RA MEDLINE=20400502; PubMed=10837471;
RA Higgins J.M.H., Cernadas M., Tan K., Irie A., Wang J.-H., Takada Y.,
RA Brenner M.B.;
RT "The role of alpha and beta chains in ligand recognition by beta 7
RT integrins.";
RL J. Biol. Chem. 275:25652-25664(2000).
CC -!- FUNCTION: INTEGRIN ALPHA-E/BETA-7 IS A RECEPTOR FOR E-CADHERIN. IT
CC MEDIATES ADHESION OF INTRA-EPITHELIAL T-LYMPHOCYTES TO EPITHELIAL
CC CELL MONOLAYERS.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA
CC SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAINS LINKED BY A
CC DISULFIDE BOND. ALPHA-E ASSOCIATES WITH BETA-7.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: EXPRESSED ON A SUBCLASS OF T-LYMPHOCYTES KNOWN
CC AS INTRA-EPITHELIAL LYMPHOCYTES WHICH ARE LOCATED BETWEEN MUCOSAL
CC EPITHELIAL CELLS.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VMFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC -!- SIMILARITY: Contains 1 VMFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD103 entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd103.htm".

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).

CC EMBL; L25851; AAB59359.2; --
CC EMBL; AF168787; AAR43107.1; --
CC PIR; A53213; A53213.
CC HSP; P11215; IABX.
CC GENE; HGNC:6147; ITGAE.
CC MIM; 604682; --
CC GO; GO:0008305; C: integrin complex; TAS.
CC GO; GO:0004895; F: cell adhesion receptor activity; TAS.
CC InterPro; IPR00413; Integrin_alpha.
CC InterPro; IPR002035; VWF_A.
CC Pfam; PF01839; FG-GAP; 3.
CC Pfam; PF00357; Integrin_A; 1.
CC PRINTS; PR00092; vwa; 1.
CC PRINTS; PR01185; INTEGRINA.
CC PRINTS; PR00453; VWFADOMAIN.
CC SMART; SM00191; Int_alpha; 3.
CC SMART; SM00327; VWA_1.
CC PROSITE; PS00242; INTEGRIN_ALPHA; 1.
CC PROSITE; PS0234; VWF_A; 1.
CC Signal; Repeat; Polymorphism; Magnesium; Calcium.
CC SIGNAL 1 18
CC CHAIN 19 1179 INTEGRIN ALPHA-E.
CC CHAIN 19 177 INTEGRIN ALPHA-E LIGHT CHAIN.
CC CHAIN 179 1177 INTEGRIN ALPHA-E HEAVY CHAIN.
CC DOMAIN 19 1124 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 1125 1147 POTENTIAL.
CC DOMAIN 1148 1179 CYTOPLASMIC (POTENTIAL).
CC DOMAIN 181 198 GLU-RICH (ACIDIC).
CC REPEAT 1 198 FG-GAP 1.
CC REPEAT 2 198 FG-GAP 2.
CC REPEAT 3 198 X-DOMAIN (EXTRA DOMAIN).
CC DOMAIN 145 199 VWF_A.
CC DOMAIN 200 391 FG-GAP 3.
CC REPEAT 401 456 FG-GAP 4.
CC REPEAT 457 506 FG-GAP 5.
CC REPEAT 510 571 FG-GAP 6.
CC REPEAT 573 638 FG-GAP 6.

FT	REPEAT	641	693	FG-GAP 7.
FT	CA_BIND	522	530	POTENTIAL.
FT	CA_BIND	586	594	POTENTIAL.
FT	CA_BIND	654	662	POTENTIAL.
FT	CA_BIND	1150	1154	GPKR MOTIF.
FT	DISULFID	70	79	BY SIMILARITY.
FT	DISULFID	126	159	BY SIMILARITY.
FT	DISULFID	706	762	BY SIMILARITY.
FT	DISULFID	823	829	BY SIMILARITY.
FT	DISULFID	893	907	BY SIMILARITY.
FT	DISULFID	1008	1033	BY SIMILARITY.
FT	DISULFID	1041	1057	BY SIMILARITY.
FT	CARBOHYD	49	49	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	271	271	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	321	321	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	444	444	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	726	726	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	782	782	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	857	857	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	934	934	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	954	954	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1065	1065	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1096	1096	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	VARIANT	360	360	D -> E.
FT	VARIANT	1041	1041	/FTID=VAR_008884.
FT	MUTAGEN	208	208	/FTID=VAR_008885.
FT	MUTAGEN	316	316	D->A: LOSS OF E-CADHERIN BINDING.
FT	CONFLICT	477	477	V -> I (IN REF. 3).
FT	CONFLICT	482	482	Q -> R (IN REF. 3).
FT	CONFLICT	950	950	R -> W (IN REF. 3).
FT	CONFLICT	1019	1019	A -> V (IN REF. 3).
FT	SEQUENCE	1179	1179	AA; 130088 MW; ES58902EDF9D95E1 CRC64;
Qy	5	VLLLTALTLC	CHGFLD	TENANTFOENARGFGQSVOLQG-----SRVVCARQEIIVAAQ 59
Db	7	LLCIASLAL	LAAFNV	DVARPMLTPKGGAPVLSLLHQDSTNQTLLVSPR---TKRT 63
Qy	60	RGLYQCDY	STGS--	CEPI-RLQVP-----VBAVNMSLGSLAATTPPQLACGPTVH 110
Db	64	PGPLHRC	SLVQDE	ILCHPVEHVPFKGRHRTVTVSRSHGVLCI-----QVLVRRP--H 116
Qy	111	QTCSENTY	VKGLCL	FLGSLNRQPO-----135
Db	117	SLSSELT--	GTCSLL	GPDLRPOAQNFFDLENLLDPDARVDTGDCYSNKEGGEDDVT 173
Qy	136	-KPEALRG	CFQED--	-----SDIAFLDGSGLIIPHDFRMKEFVSTVMEQL- 180
Db	174	ARQRALE	KEEEDKE	EBEDEEBEAEAGTEAIIIDGSGSIDPDPQAKDFISNMENFY 233
Qy	181	-KSKNTL	FLSLMY	SEFRHIFTPEFQNNPNPSLVKPTOLLGRTHATGIRKVVRELF 239
Db	234	EKCFCNF	ALVQYGV	IQTEFDLRSQDVMAVLARVQNTITQVGSVTKTASAMQHVLDISF 293
Qy	240	NITNGARK	NAFKIL	VITDGEKEDPLGVEDVITPEADREGVIRYVLGVGDFAFSEKROE 299
Db	294	TSSGSR	RKASKV	MVITDGGIFEPLNTTVINSFMQGVVERFALGVGEFFSARTARE 353
Qy	300	LNTIAS	KPPRDH	VOVNFPEALTIQNLREKIFAIEGTQTGSSSFHEHMSDEGSAAI 359
Db	354	LNLIAS	DPDETH	AFKVTNYMALDGLSLKLYNIISMEGT---VGDALHYQLAQIGFSAQI 410
Qy	360	TSNGP-	LLSTVGS	YDWAGVFLY-TSKEKSTFTNMTFRVDSKXDA---YLGVAAILLR 413
Db	411	LDERQV	LLGAVG	AFDWSGGALLYDTSRRGRFLNQFAAAADAAQAQSYLVGFAVLEK 470
Qy	414	NRVQSL	VLGAPRY	QHIGLVAMFR-QNTGWMESNAVY-KGTQICAYFCASICSDVDISNGS 471

Query Match 19.5%; Score 1163; DB 1; Length 1179;
Best Local Similarity 28.8%; Pred. No. 2.7e-69;
Matches 356; Conservative 221; Mismatches 481; Indels 180; Gaps 41;

FT DISULFID 998 1023 BY SIMILARITY.
PT DISULFID 1031 1047 BY SIMILARITY.
FT CARBOHYD 51 51 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 256 256 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 314 314 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 341 341 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 364 364 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 418 418 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 437 437 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 718 718 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 773 773 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 823 823 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 846 846 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 911 911 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 925 925 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 968 968 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1013 1013 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1055 1055 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1086 1086 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1167 AA; 128983 MW; B83331C115DCCCFD CRC64;
Query Match 19.6%; Score 1165.5; DB 1; Length 1167;
Best Local Similarity 29.1%; Pred. No. 1.8e-69;
Matches 362; Conservative 213; Mismatches 465; Indels 205; Gaps 43;
QY 5 VLLTALITLCHGFNLDTEA--MTQENARGFGQSVQLOGSRVVGAPQELVAANORGS 62
DB 8 LLCMASLKPQGFANMDVMAWTALQPGAPAVLSLLHLDPEN-----NOTCLLVARSS 62
QY 63 -----LVQCDSTGSCSIRIQVVEAVNMSELGLSLAATT--SPQLLAC-GPTVHQT 113
DB 63 NENTAALYRCALSI-SPDEIAQ-PVEHICMPKRGYQVTLVNGHENGVLVCIOQAKFR 120
QY 114 SENTVVGKLCFTFGNLRQOPKPEALRG-----C----- 144
DB 121 SUNSELTGACSLTFLNLDLQAOYFSDLEGFLDPGARVDSGDCSKSGTGECKSARR 180
QY 145 -----PODSIDIAFLIDGSGSIIPIHDFRMKEFVSTVMEQL--KKSITLFLMOYS 193
DB 181 RRTVBEDEEETGTAIVLDGSGSIGPSDFQKAKNFISTMRNFYKCFECFALVOYG 240
QY 194 EEFRIHFTKPEFQNNPNSLVKPTTOLLGRTHATGINKVRELFNITNGARKNAFKIL 253
DB 241 AVIOTFDFLQESRDINASLAKVQSVQVKEVTKTASAMQHVLDNITFIPSGRSKKALKVM 300
QY 254 VVITGKEGDPGLGYEDVTPADREGVIRYVVGVDAPSEKSRQBLATIASKPPRDVFE 313
DB 301 VVITDGDIFGDFLNTLVINSPEMQGVRFALGVDRFANNYRELKLLIASDPKEATP 360
QY 314 QVNNPEALKTIONQLREKIFAJEGTGTGSSSFHEMSEQGSAAITNSGP-LLSTVGSY 372
DB 361 KVTNYSALDGLLSKLQORIVHMEGT---VGDALQYLAQTGFSQAQILDKGVLLGTVGAF 417
QY 373 DWAGGVFLY-TSKERSTENMT-RVDS-DMNDAYLGVAATILNRVQSLVLAGAPYQHI 429
DB 418 NMSGGALLYSTONGRCFLNQTAKEDSRVQSYLSGLSLAVLHKAGISYVAGAPRHKUR 477
QY 430 GLVAFRQNTGHWESNA---NVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYRQT 486
DB 478 GAVFELRKEDR-BEDAPVRRIEGEQMGSYFGSVLCPVDIDMDGTTDFLLVAAPFYHIG 535
QY 487 RCGQVSVCLPLPGQARWOCDAVLGEQCPMGRGCAALTVLGDVNGDKLTDAVAGAP-- 544
DB 536 BEGRVYVQVPE-QDASFSLAHLTSGHPGLTMSRFGFAMAAVGDINDQKFTDVAIGAPLE 594
QY 545 -----GREDNRGAVYLFHGTGSGSISPSHSQRIAGSKLSPLQYFGQSLSGGQDLTMDGLV 600
DB 595 GFGAGGASVGSYVLYNGHSG-GLVDSFSGQIRASSVASGLHYFGMSVSGGLDFNGDGLA 653
QY 601 DLTVAGQGVHLLRSQPLRVKALMEFNPREVARNVFCNDQVYVKGKAGEVRVCLHVQK 660
DB 654 DITVGRSDSAVLRGRPVLDLTVSMFTFP-----DALEPMVFTGKM--DYNLCFEVDS 703

QY 661 S---TRDLREGQIQSVVTVYDLALDSCRPSPRAVFNETKNSTRQTVGLGATYC----- 712
DB 704 SVVASBEGLAEMFLNFTVDVV-----TKQORLQCEDSSGCGCLRKWN 748
QY 713 -----ETLKLQENCIEDPVSPIVLNLSLVCGLPSAENLR---PVLAE 756
DB 749 GGSFLCBFWLISFEEL-----CEDCFENITIKVYF-----FQTSGRDRDYNPTL--D 797
QY 757 AQRLLFTALP--PPEKNQNDNIQDDLSITFSFMSLDCLVVGGRPFNFVTVVRNDS 814
DB 798 HYKEPSAIFQLPYEKDCKNKVFCIAEQLTN-ISOQELVVGTVKEVTMNLSTNSGDS 856
QY 815 YRTQVTFPPFLDLSYRKVSTLQNSORSORSLACESASSTEVSGALKSTSCSINHPIFE 874
DB 857 YMTNMAIYPRNLQFKKI-----QKPVSDVQCDPKPV---ASVLWMCKIGHPL-K 906
QY 875 NSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEFQLELPKVAVMVVTSHG 934
DB 907 RSSVNVSVTWQLESVFPNETADITVTSNKSLEARETR---SLQFRAFIATLSR-- 961
QY 935 VSTKYLNTASNTSRVHQYQVSNLQORSPLISLVFLVPLNQTIVWDRPOVTSEN 994
DB 962 PSVMYAN--TSQSPSDHKEFFPNVHGENLFGAVFQLQICVPIKLQDF-----QIVRVKN 1013
QY 995 LSST-----CHTKERLPSHDSFLAELKAPVWNCIAVCQICQIDIFFFGIOBEFNATLK 1049
DB 1014 LTKQDTHETQSGEPACGSDPVQHVKEHVSVCAL-----TSNK 1053
QY 1050 GNLSFDWIKYKSNHLLIVSTA-----EILFNDSVFTLLPGQAFVRSOTETKVEFF-- 1101
DB 1054 ENVTVAEISVGHTKQLLRDVSBLPILGEISFNKSLYEGLNAE-----NHRTKITVIFL 1107
QY 1102 --EVPNPLPLVGSVGLLALLALITAAALYKLGFFKROVKDMSE 1144
DB 1108 KEETRSPLIIGSIGLLVWVILFALFCGFFKRIQOENLE 1152
RESULT 8
ITAE HUMAN
ID ITAE HUMAN STANDARD; PRT; 1179 AA.
AC P38570; QNZU9;
DT 01-OCT-1994 (Rel. 30, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-E precursor (Mucosal lymphocyte-1 antigen) (HML-1
DE antigen) (CD103 antigen) (Integrin alpha-IEL).
GN ITGAE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 19-38 AND 179-188.
RC TISSUE=Leukemia, and Lymphocytes;
EX MEDLINE=94164962; PubMed=8119947;
RA Shaw S.K.; Cepek K.B.; Murphy E.A.; Russell G.J.; Brenner M.B.,
RA Parker C.M.;
RT "Molecular cloning of the human mucosal lymphocyte integrin alpha E
RT subunit. Unusual structure and restricted RNA distribution."
RL J. Biol. Chem. 269:6016-6025(1994).
RN [2]
RP REVISIONS TO 88-114.
RA Parker C.M.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 53-1179 FROM N.A.
RC TISSUE=Fetal kidney;
EX MEDLINE=20138496; PubMed=10673275;
RA Touchman J.W.; Anikster Y.; Dietrich N.L.; Maduro V.V.; McDowell G.,
RA Shotelersuk V.; Bouffard G.G.; Beckstrom-Sternberg S.M.; Gahl W.A.,
RA Green E.D.;
RT "The genomic region encompassing the nephropathic cystinosis gene
RT (CTNS): complete sequencing of a 200-kb segment and discovery of a

Db 342 SFNLSGSGISADLSKSHAVGVAGKAGGFGFLDREDLQAGTFVQCBPLTSDVRGGY 401
Qy 404 LGYAAA-IILRNRYQSLVGLGAPRYOHICLVAMPR--ONTGWESNANVKGFTQICAYGAS 460
Db 402 LGYVAMWTSRSSRPLLAAGAPRYOHVGVLLFOAPAGGRWNTQIKGFTQICSGYFGE 461
Qy 461 LCSVDVDSNGSTDLVLGAPHYYQTRGGQVSVCLPRGQBARWCDVAVLYGEOGPWGR 520
Db 462 LCSVDLDQDGRABLLICAPLFFGQGRGVFTY---QRRSLFEMVSELDQDPGYPLGR 518
Qy 521 FGAALTVLGVNDKLTDAVAGPGRDNRCNAVLFECTSGSISPSHSORLAGSKLSR 580
Db 519 FGAALTALTDINGRLTDVAGAPLEB--QNAVTFNGKXP-GUSPQPSORIOGAQVFP 575
Qy 581 LOYFQSLSGQDUTMDGLVDLTVGAGHVLNLSQVPLRVKALMBFNPREVARNVFECN 640
Db 576 IRWEGRIHGVKDLGGDLADVVVGAEGRVVLSRFDVVTLSLSPFVHEVECS 635
Qy 641 DOVVKGEAG-EVAVCLFVQKSTDRRLREGQIOSVVTYDLADSGRPHSAVFNETKST 699
Db 636 YSAREEQHGVKACAFRIKPLTPQ--FQGRLLANLSTYLDGHRMRSLGFLPDGSHL 693
Qy 700 RROTQVLGLTQCTBTLKQLQPNCTEDPVSPIVLKNFSLV---GTPLSAFEN-LRPVLAE 755
Db 694 SGNSTIIP-DKSCLDHFHPICIQDLISPINVSLNLSLEEBCTPRDQGRAMQPILRP 752
Qy 756 DAQRLFTALPFKXKCNNDNICQDLSITFSFMSLDCLVGGP-----REFNVTVTVRN 809
Db 753 SIHTV-TKEIPFKNCGEDKCEANLTLSSPARS-----GPLRLMSSASLAVETLSN 804
Qy 810 DGEDSYRTQVTFPPDLVSRKYSTLQNRQSRWRLACESASSTEVSGAL-KSTSCIN 868
Db 805 SGEDAYVRLDLDFRGLSFRKVMQ---PHSRMPVSCHEL--TEGSSLLTKLKNSV 859
Qy 869 HPIPPENSEVNTITFDVDSKASIGNKLLKANVTSEN-NMPTNKTEFLELPVKYAVY 927
Db 860 SPIFKAQEVSLQWMTLLNSEDVFNLTGTVHCENSSLSQEDNSAATHIPVLYPVN 919
Qy 928 MVTSHGVSTKYNLFTASENTSRVQHOYOVNLSQESLPSLVFLVPVNLQTVIWDPR 987
Db 920 ILTKEQENSTLYISFTPKGPKTQVQVYQV----RIQSAVDHNMPT-LEALVGVPRP 973
Qy 988 Q-----VTFSENLS--TCHTKE-RLPSSHDFLAEALRKAAPVNCSTIAVQCQICDIPFF 1038
Db 974 HSEDLIYTVSVQTDPLVTCHEBCLKRPSSE---AEQFCILPV-----QFRCPIVE- 1021
Qy 1039 GIQEEFNATLKNLSFDWYIKTSHNLLIYSTAEILEPNDVSFTLLPQAGAFVRSQETKV 1098
Db 1022 --RWEHLIQTGVFELSKIRAS-STLSLCSLSVSFNSKHFHYGSKA-SBAQVLVKV 1077
Qy 1099 EPFEVPNPLPLVSSVGGLLALLALITAAALYKLGFFKRYQKDMW-SGGSPGGAEP 1152
Db 1078 DLIHEKEMLVVVLGGGLVLLFLIPLALYKVGFFKRNKRXEADGGVPNSGP 1132
RESULT 7
ID ITAE MOUSE STANDARD; PRT; 1167 AA.
AC Q60677;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-E precursor (Integrin alpha M290).
GN ITGAE.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AKR;
RX MEDLINE=95187992; PubMed=7882170;

RA Smith T.J., Ducharme L.A., Shaw S.K., Parker C.M., Brenner M.B.,
RA Kilshaw P.J., Weis J.H.;
RT "Murine M290 integrin expression modulated by mast cell activation.";
RL Immunity 1:393-403 (1994).
CC -1- FUNCTION: INTEGRIN ALPHA-E/BETA-7 IS A RECEPTOR FOR E-CADHERIN. IT
CC MEDIATES ADHESION OF INTRA-EPITHELIAL T-LYMPHOCYTES TO EPITHELIAL
CC CELL MONOLAYERS. MICE EXPRESSING A NULL MUTATION OF THE ALPHA-E
CC SUBUNIT GENE EXHIBIT A MARKED REDUCTION IN THE NUMBERS OF
CC INTRAEPITHELIAL LYMPHOCYTES IN THE GUT AND IN THE DEVELOPMENT OF
CC GUT-ASSOCIATED LYMPHOCYTE AGGREGATES, SUPPORTING A SPECIFIC ROLE FOR
CC THIS INTEGRIN IN MEDIATING RETENTION OF LYMPHOCYTES IN THE
CC INTESTINAL WALL.
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA
CC SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAINS LINKED BY A
CC DISULFIDE BOND. ALPHA-E ASSOCIATES WITH BETA-7.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -1- SIMILARITY: Belongs to the integrin alpha chain family.
CC -1- SIMILARITY: Contains 1 VWFA domain.
CC -1- SIMILARITY: Contains 7 FG-GAP repeats.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U12236; AAC52142.1; -;
DR HSSP; P11215; 1A8X.
DR MGD; MGI:1298377; Itgae.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR02035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01145; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 3.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWFA; 1.
DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Signal; Repeat; Magnesium;
KW Calcium.
FT SIGNAL 1 19 BY SIMILARITY.
FT CHAIN 20 1167 INTEGRIN ALPHA-E.
FT CHAIN 20 181 INTEGRIN ALPHA-E LIGHT CHAIN.
FT CHAIN 183 1167 INTEGRIN ALPHA-E HEAVY CHAIN.
FT DOMAIN 20 1114 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1115 1137 POTENTIAL.
FT DOMAIN 1138 1167 CYTOPLASMIC (POTENTIAL).
FT REPEAT ? ? FG-GAP 1.
FT REPEAT ? ? FG-GAP 2.
FT DOMAIN 149 192 X-DOMAIN (EXTRA DOMAIN).
FT DOMAIN 193 384 VWFA.
FT REPEAT 449 501 FG-GAP 3.
FT REPEAT 503 584 FG-GAP 4.
FT REPEAT 566 631 FG-GAP 5.
FT REPEAT 634 686 FG-GAP 6.
FT REPEAT 634 686 FG-GAP 7.
FT CA_BIND 514 522 POTENTIAL.
FT CA_BIND 578 586 POTENTIAL.
FT CA_BIND 646 654 POTENTIAL.
FT DOMAIN 185 191 GLU-RICH (ACIDIC).
FT SITE 1140 1144 GPFKR MOTIF.
FT DISULFID 72 83 BY SIMILARITY.
FT DISULFID 130 164 BY SIMILARITY.
FT DISULFID 698 754 BY SIMILARITY.
FT DISULFID 814 820 BY SIMILARITY.
FT DISULFID 884 898 BY SIMILARITY.

Db 1083 VYKQMLYLVLSIGIGLLILLIFVLYKVGFKRLKEMRAGRCVPGNIP 1136

RESULT 6
ID ITAL_MOUSE STANDARD; PRT; 1163 AA.
AC P24063;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-L precursor (Leukocyte adhesion glycoprotein LFA-1
alpha chain) (Leukocyte function associated molecule 1, alpha chain)
(CD11a).
GN ITGAL OR LFA-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91268576; PubMed=2051027;
RA Kaufmann Y, Tseng E., Springer T.A.;
RT "Cloning of the murine lymphocyte function-associated molecule-1
alpha-subunit and its expression in COS cells.";
RL J. Immunol. 147:369-374 (1991).
RN [2]
RP SEQUENCE OF 24-42.
RX MEDLINE=8518276; PubMed=3887182;
RA Springer T.A., Toplow D.B., Dreyer W.J.;
RT "Sequence homology of the LFA-1 and Mac-1 leukocyte adhesion
glycoproteins and unexpected relation to leukocyte interferon.";
RL Nature 314:540-542 (1985).
CC -1- FUNCTION: INTEGRIN ALPHA-L/BETA-2 IS A RECEPTOR FOR ICAM1, ICAM2,
ICAM3 AND ICAM4. IS INVOLVED IN A VARIETY OF IMMUNE PHENOMENA
INCLUDING LEUKOCYTE-ENDOTHELIAL CELL INTERACTION, CYTOTOXIC T-CELL
MEDIATED KILLING, AND ANTIBODY DEPENDENT KILLING BY GRANULOCYTES
AND MONOCYTES. MICE EXPRESSING A NULL MUTATION OF THE ALPHA-L
SUBUNIT GENE DEMONSTRATE IMPAIRED TUMOR REJECTION AND IMPAIRED
LEUKOCYTES RECRUITMENT.
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-L
ASSOCIATES WITH BETA-2.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: LEUKOCYTES.
CC -1- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -1- SIMILARITY: Belongs to the integrin alpha chain family.
CC -1- SIMILARITY: Contains 1 VWFA domain.
CC -1- SIMILARITY: Contains 7 FG-GAP repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M60778; AAA39426.1; -.
DR PIR; I56126; I56126.
DR HSSP; P20701; 1LFA.
DR MGD; MGI:96606; Itgal.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR SMART; SM00453; VWFADOMAIN.
DR SMART; SM00191; Int alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS02324; VWFA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;

Signal; Magnesium; Calcium;

Repeat.
KW SIGNAL 1 23
FT CHAIN 24 1163
FT DOMAIN 24 1084
FT TRANSMEM 1085 1108
FT DOMAIN 1109 1163
FT REPEAT 39 88
FT REPEAT 148 334
FT DOMAIN 399 454
FT REPEAT 455 514
FT REPEAT 516 573
FT REPEAT 576 628
FT CB_BIND 466 474
FT CA_BIND 528 536
FT CA_BIND 588 596
FT SITE 1111 1115
FT DISULFID 70 77
FT DISULFID 108 126
FT DISULFID 147 199
FT DISULFID 651 703
FT DISULFID 767 773
FT DISULFID 840 856
FT DISULFID 993 1009
FT DISULFID 1017 1048
FT CARBOHYD 86 86
FT CARBOHYD 185 185
FT CARBOHYD 270 270
FT CARBOHYD 444 444
FT CARBOHYD 668 668
FT CARBOHYD 696 696
FT CARBOHYD 724 724
FT CARBOHYD 728 728
FT CARBOHYD 776 776
FT CARBOHYD 857 857
FT CARBOHYD 880 880
FT CARBOHYD 890 890
FT CARBOHYD 899 899
FT CARBOHYD 927 927
FT CARBOHYD 1056 1056
SQ SEQUENCE 1163 AA; 128343 MW; A7A3078489B8232F CRC64;

Query Match 26.0%; Score 1549; DB 1; Length 1163;

Best local Similarity 34.2%; Pred. No. 5e-95;

Matches 409; Conservative 216; Mismatches 464; Indels 106; Gaps 38;

Qy 1 MALRV---LLLTALT---CHGFNLDTENAMTFQENA-RGFQGVVQLGSRVWVGAPQ 52
Db 1 MSFRIAGPRLLILGLQLPAKMSYNLDTRTQSLAQAGHFQYVQLIEDG-VVVGAPG 59
Qy 53 EIVAAHQSGLYOCYDSTGSCBPIRLQVPEAVNMSIGLSLAATTPPQLLAGCPVHOT 112
Db 60 E---GDNLTGLYHCRTSSEFCQPSVLSH-GSNHNTSKYLGMTLATDAKAGSLACDPLGSR 115
Qy 113 CSNFTYVKGICFLFGSNLRQPOKEPEALRGCPQEDSDIAFLIDGSGSIIPHPFRMKEF 172
Db 116 CDQNTYLSGLCYLFPQSLGPMQNPAYQECMKGVLDVLFDDSGSLDRKFEKILEP 175
Qy 173 VSTVMEQLKSKTLFSLMAYSEBFIHFTPEK-ONNPNSRSLVKPITQLGRHTHTATGI 231
Db 176 MKDVWRKLSNTSYQAAVQFSTDCRTEFTLDYVVKQKNSDVLGSGVQPMFLNTTPRAI 235
Qy 232 RKVRELFNITNGARKNAFKILVITDGEKFGDPLGYEDVIPLEADREG-----VIRYV 284
Db 236 NYVAVHVFKEBSGARPDATKVLVITD-----EASDKGNISAAHDITRYI 281
Qy 285 IGVEDAFRSBKSQELNTIASKPPRDHVFOVNNFEALKTIONQLREKIFAEIGTQSSS 344
Db 282 IGIGKHFVSVQKTKTLHI PASEPVEBPKLDTFEKLDKLTDLQRIYAEIGNRDLT 341
Qy 345 SFEHMSQEGFSAAITNSGPNLLSTVGSYDWAGGVF-LYTSKEKSTFINMTVEVSDMNDAY 403


```

Db 901 ENKASSKATQLELPVKIATYTWISROESTKTFNFATSEKOKKAEHRIRVNNLSQ 960
QY 964 RSLPISLVFLVRLNQTVINDRPQVTFSENLSSCTKRLPSHSDFLAELRKAPVYVC 1023
Db 961 RLALISINFWPVLNGVAVMDVMEAPSQL--PCVSRKPPQSHSDFLTQISRSPMLDC 1018
QY 1024 STAVCORLOCDPFGICGEEFNATLGNLSFDWYIKTSHNLLIYSTAEIILFNDSVFTLL 1083
Db 1019 STADCLQRCDFVPSVGEELDTLGNLSFGVRETLQKRVLVVSVARITFDTSVISQL 1078
QY 1084 PQGAFVRSQETKVEPEVEPNFLVGVSSVGGLLALITAAALYKLGFFFRQYKQMS 1143
Db 1079 PQEAFMAQXEMVLEEDVEVYNAIFLMGSSVGLLALITATLYKLGFFFRKRYKEMLE 1138
QY 1144 E 1144
Db 1139 D 1139

RESULT 5
ID _ITAL_HUMAN STANDARD; PRT; 1170 AA.
AC P20701; OX3746;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Integrin alpha-L precursor (Leukocyte adhesion glycoprotein LFA-1
DE alpha chain) (Leukocyte function associated molecule 1, alpha chain)
DE (CD11a).
GN ITGAL OR CD11A.
OS Homo sapiens (Human).
OC Sukariyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1), AND PARTIAL SEQUENCE.
RA MEDLINE=89139587; PubMed=2537322;
RA Larson R.S., Corbi A.L., Berman L., Springer T.;
RT "Primary structure of the leukocyte function-associated molecule-1
RT alpha subunit: an integrin with an embedded domain defining a protein
RT superfamily.";
RL J. Cell Biol. 108:703-712(1989).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RX MEDLINE=99425270; PubMed=10493829;
RA Loftus B.J., Kim U.-J., Sneddon V.P., Kalush F., Brandon R.,
RA Fuhrmann J., Mason T., Crosby M.L., Barnstead M., Cronin L.,
RA Deslattes Mays A., Cao Y., Xu R.X., Kang H.-L., Mitchell S.,
RA Eichler E.E., Harris P.C., Venter J.C., Adams M.D.;
RT "Genome duplications and other features in 12 Mb of DNA sequence from
RT human chromosome 16p and 16q.";
RL Genomics 60:295-308(1999).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 153-335, AND REVISION TO 214.
RX MEDLINE=96036067; PubMed=7479767;
RA Qu A., Leahy D.J.;
RT "Crystal structure of the I-domain from the CD11a/CD18 (LFA-1, alpha
RT L beta 2) integrin.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:10277-10281(1995).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 153-335.
RX MEDLINE=96398682; PubMed=8805579;
RA Qu A., Leahy D.J.;
RT "The role of the divalent cation in the structure of the I domain
RT from the CD11a/CD18 integrin.";
RL Structure 4:931-942(1996).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 153-334.
RX MEDLINE=99425288; PubMed=10493852;
RA Kallen J., Welzenbach K., Ramage P., Geyl D., Kriwacki R., Legge G.,
RA Cottens S., Weitz-Schmidt G., Hommel U.;
RT "Structural basis for LFA-1 inhibition uponlovastatin binding to the

```

```

RT CD11a I-domain.";
RL J. Mol. Biol. 292:1-9(1999).
CC -!- FUNCTION: INTEGRIN ALPHA-L/BETA-2 IS A RECEPTOR FOR ICAM1, ICAM2,
CC ICAM3 AND ICAM4. IT IS INVOLVED IN A VARIETY OF IMMUNE PHENOMENA
CC INCLUDING LEUKOCYTE-ENDOTHELIAL CELL INTERACTION, CYTOTOXIC T-CELL
CC MEDIATED KILLING, AND ANTIBODY DEPENDENT KILLING BY GRANULOCYTES
CC AND MONOCYTES.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-L
CC ASSOCIATES WITH BETA-2.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=P20701-1; Sequence=Displayed;
CC Name=2;
CC IsoId=P20701-2; Sequence=VSP_002738;
CC Note=No experimental confirmation available;
CC -!- TISSUE SPECIFICITY: LEUKOCYTES.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -!- DATABASE: NMB=PROW; NOTE=CD guide CD11a entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd11a.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Y00796; CAA68747.1; --
DR EMBL: AC002310; AAC31672.1; --
DR PIR: S03308; S03308.
DR PDB: 1LFA; 29-JAN-96.
DR PDB: 1ZON; 07-DEC-96.
DR PDB: 1ZOO; 07-DEC-96.
DR PDB: 1ZOP; 07-DEC-96.
DR PDB: 1COP; 07-AUG-00.
DR PDB: 1DGO; 03-FEB-00.
DR PDB: 1MJN; 28-JAN-03.
DR PDB: 1MQ8; 14-JAN-03.
DR PDB: 1MQ9; 14-JAN-03.
DR PDB: 1MQA; 14-JAN-03.
DR Genew; HGNC:6148; ITGAL.
DR MTM; 153370; --
DR GO; GO:0008305; C.integrin complex; TAS.
DR GO; GO:0006928; P.cell motility; TAS.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00032; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWFA; 1.
DR Signal; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
DR Signal; 3D-structure; Magnesium; Calcium; Repeat;
KW Alternative splicing.
FT SIGNAL 1 25
FT CHAIN 26 1170
FT DOMAIN 26 1088
FT TRANSMEM 1089 1112
FT DOMAIN 1113 1170
FT REPEAT 42 91
FT REPEAT 92 149
FT DOMAIN 170 349

```


Db 8 LLLFTALATSLGNMLDTEBELTAFRVDSSAGSGDSVVOYANSVVVGAPOKITAANQTGGLY 67
QY 65 QDYSTSGCEPIRLVPPVEAVNMSLGLSLAATSPPOLLAGCGTTHQTCSENTYVKGLCF 124
Db 68 QCGYSTGACEPIGLVPPVEAVNMSLGLSLAATSPPOLLAGCGTTHQTCSENTYVKGLCF 127
QY 125 LFGSNLRQOPOKEPPEALRGCPQEDSDIAFLIDGSGSIIPEDFRMKEFVSTWMEQLKSK 184
Db 128 LLGPT--QLTQRLPVSQCEPQEQDIIVLIDGSGSISSENFATWVFNVISQFQRS 185
QY 185 TPLSMOYSEBFIHTTKEFQNNPNRSLVKEITQLLGRHTATGIRKVVRELFINITNG 244
Db 186 TOPSLMOYSEBFIHTTKEFQNNPNRSLVKEITQLLGRHTATGIRKVVRELFINITNG 245
QY 245 ARKNAKILVITDGEKFGDPLGYEDVIEADREGVIRYVIGVGDFAFSEKSEKSOELNTIA 304
Db 246 ARDATKILVITDGEKFGDPLGYEDVIEADREGVIRYVIGVGDFAFSEKSEKSOELNTIA 305
QY 305 SKPPRDHVQVNNFEALKTIONLREKIPAEIGTQGTGSSSSPEHEMSQSGFSAALITSGP 364
Db 306 SKPPRDHVQVNNFEALKTIONLREKIPAEIGTQGTGSSSSPEHEMSQSGFSAALITSGP 365
QY 365 LLSTVGSVWAGSVFLYTSKESKSTFNNTRVDSNDMDAVLYGAAAILLNRVOSLVIGAP 424
Db 366 VLGAVGSFTWGAFLYPPNMPSTFNNMQEVDNRDYSYLGYSTELALWGVQSVLIGAP 425
QY 425 RYCHIGLVAMFRQNTGMESENANVKTQICAYFGASLCSVDVDSNGSTDVLVIGAPHYVE 484
Db 426 RYCHIGLVAMFRQNTGMESENANVKTQICAYFGASLCSVDVDSNGSTDVLVIGAPHYVE 485
QY 485 QTRGGGVSCPLPRGORARWCDVLYGQCPWGRFGAALTVLGVGDKLTDVAIGAP 544
Db 486 QTRGGGVSCPLPRGORARWCDVLYGQCPWGRFGAALTVLGVGDKLTDVAIGAP 544
QY 545 GEEDNRGAVYLFHGTGSGSISPSHSQRTAGSKLSPRLQVFCQSLGCGDGLTMDGLVDITV 604
Db 545 GEEDNRGAVYLFHGTGSGSISPSHSQRTAGSKLSPRLQVFCQSLGCGDGLTMDGLVDITV 604
QY 605 GAQHVLLRSPVLRVKAIMEFNPREVARNVFEQNDQVQKKEAGEVRVCLVHVKSTRD 664
Db 605 GARGQVLLRTPVLRVKAIMEFNPREVARNVFEQNDQVQKKEAGEVRVCLVHVKSTRD 664
QY 665 RLREGIOQVWVYDIALDSGRPHSAVFNETKSTROTQVLTGLTQTCETLKLQLPNCIE 724
Db 665 RLREGIOQVWVYDIALDSGRPHSAVFNETKSTROTQVLTGLTQTCETLKLQLPNCIE 724
QY 725 DVPSPVLRNLPSLVGTPLSAGNLRPVLAEDAQRLTALPPEKXCGNDICQDLSIT 784
Db 725 DVPSPVLRNLPSLVGTPLSAGNLRPVLAEDAQRLTALPPEKXCGNDICQDLSIT 784
QY 785 RFSMSLDCLVVGPREENVTVVRNDGDSYRTQVTFPFLDLSYRKYSTLQNESQSW 844
Db 785 RFSMSLDCLVVGPREENVTVVRNDGDSYRTQVTFPFLDLSYRKYSTLQNESQSW 844
QY 845 RIACSSASTVSGALKSTCSINHPITPENSEVTNITFDVDSKASLGNKLLKANVTS 904
Db 845 RIACSSASTVSGALKSTCSINHPITPENSEVTNITFDVDSKASLGNKLLKANVTS 904
QY 905 ENNMERTKTEPQELPKVAYVMTVTSHTGYSKYLNTAS-ENTSRVMOHQYQVSNLQ 963
Db 905 ENNMERTKTEPQELPKVAYVMTVTSHTGYSKYLNTAS-ENTSRVMOHQYQVSNLQ 963
QY 964 RSLPILSVLVPVRLNQTIVINDRPQVTFPENSELSTCHTKERLPSHDSFLAELRAPVNC 1023
Db 964 RSLPILSVLVPVRLNQTIVINDRPQVTFPENSELSTCHTKERLPSHDSFLAELRAPVNC 1023
QY 1024 SIACQRTQCDIPPGIOEENFATLKNLSFDWYIKTSHNHLIVSTRAELFNDSVFTLL 1083
Db 1024 SIACQRTQCDIPPGIOEENFATLKNLSFDWYIKTSHNHLIVSTRAELFNDSVFTLL 1083
QY 1083 SIAGCLRFCDVPSFVQSEEDLFTLKNLSFGWRAQILQKQVSVVAEITFDTSVSQL 1082
Db 1083 SIAGCLRFCDVPSFVQSEEDLFTLKNLSFGWRAQILQKQVSVVAEITFDTSVSQL 1082
QY 1084 PQQAFVRSQTEKVEPEVENPLPLVGVSSVGGLLALLATLALYKLGFFKROVDMWS 1143
Db 1084 PQQAFVRSQTEKVEPEVENPLPLVGVSSVGGLLALLATLALYKLGFFKROVDMWS 1143
QY 1083 PQQAFVRSQTEKVEPEVENPLPLVGVSSVGGLLALLATLALYKLGFFKROVDMWS 1142
Db 1083 PQQAFVRSQTEKVEPEVENPLPLVGVSSVGGLLALLATLALYKLGFFKROVDMWS 1142

QY 1144 E 1144
Db 1143 E 1143

RESULT 4

ITAD HUMAN STANDARD; PRT; 1162 AA.
AC Q13349; Q15575; Q15576;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-D precursor (Leukointegrin alpha D) (CD11d) (AD82).
GN ITGAD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RX MEDLINE=96111956; PubMed=8777714;
RA Van der Vliet M., Le Trong H., Wood C.L., Moore P.F., St John T.,
RA Staunton D.E., Gallatin W.M.;
RT "A novel leukointegrin, alpha d beta 2, binds preferentially to ICAM-3";
RL J. Biol. Chem. 275:8959-8969(2000).
RN [2]
RP SEQUENCE OF 1-235 FROM N.A.
RX MEDLINE=20187620; PubMed=10722744;
RA Noti J.D., Johnson A.K., Dillon J.D.;
RT "Structural and functional characterization of the leukocyte integrin gene CD11d. Essential role of Sp1 and Sp3";
RL J. Biol. Chem. 275:8959-8969(2000).
RN [3]
RP SEQUENCE OF 467-537; 571-602; 633-666; 788-834 AND 910-1125 FROM N.A.
RX MEDLINE=96257236; PubMed=8666289;
RA Wong D.A., Davis E.M., LeBeau M., Springer T.A.;
RT "Cloning and chromosomal localization of a novel gene encoding a human beta 2-integrin alpha subunit";
RL Gene 171:291-294(1996).
RN [4]
RP INTERACTION WITH VCAM1.
RX MEDLINE=99059844; PubMed=9841932;
RA Grayson M.H., Van der Vliet M., Sterbinsky S.A., Michael Gallatin W.,
RA Hoffman P.A., Staunton D.E., Bochner B.S.;
RT "alpha beta 2 integrin is expressed on human eosinophils and functions as an alternative ligand for vascular cell adhesion molecule 1 (VCAM-1)";
RL J. Exp. Med. 188:2187-2191(1998).
RN [5]
RP INTERACTION WITH VCAM1.
RX MEDLINE=99370002; PubMed=10438935;
RA Van der Vliet M., Crowe D.T., Hoekstra D., Vazeux R., Hoffman P.A.,
RA Grayson M.H., Bochner B.S., Gallatin W.M., Staunton D.E.;
RT "The leukocyte integrin alpha D beta 2 binds VCAM-1: evidence for a binding interface between I domain and VCAM-1";
RL J. Immunol. 163:11984-11990(1999).
CC -!- FUNCTION: INTEGRIN ALPHA-D/BETA-2 IS A RECEPTOR FOR ICAM3 AND VCAM1. MAY PLAY A ROLE IN THE ATHEROSCLEROTIC PROCESS SUCH AS CLEARING LIPOPROTEINS FROM PLAQUES AND IN PHAGOCYTOSIS OF BLOOD-BORNE PATHOGENS, PARTICULATE MATTER, AND SENESECENT ERYTHROCYTES FROM THE BLOOD.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-D
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: EXPRESSED MODERATELY ON MYELOMONOCYTIC CELL LINES AND SUBSETS OF PERIPHERAL BLOOD LEUKOCYTES AND STRONGLY ON TISSUE-SPECIALIZED CELLS, INCLUDING MACROPHAGES FOAM CELLS WITHIN ATHEROSCLEROTIC PLAQUES, AND ON SPLENIC RED PULP MACROPHAGES.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.

CC	-/- SIMILARITY: Contains 1 VWFA domain.	QY	1	MALRVLLLTALTTLCHGNVLDNTENMTQENRARGQSVVQLQGSRRVVGADQEIIVANQR	60
CC	-/- SIMILARITY: Contains 7 FG-GAP repeats.	DB	1	MTLXALLVLTALACHGNVLDTEHPMTFQENAKGQGVVQLGGTSVVVAAPQEKAVNQT	60
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).	QY	61	GSLVQCDYSTGSCPIRLQVPEAVNMSLGLSLAATSPQQLACGPTVHQTCSNTYVK	120
CC		DB	61	GALVQCDYSTGRCHPIPLQVPEAVNMSLGLSLAVSTVPQQLACGPTVHQNCKENTYV	120
CC		QY	121	GLCFPLGSLNRQOQKQKPBALRGCPQEDSDIAFLIDSGSIIPHDFRMKGFVSTVMEQL	180
CC		DB	121	GLCYLFGSLNRPPQPPPEALRECFQESDITVFLIDSGSINNIDFQOKKGFVSTVMEQF	180
CC		QY	181	KKSKTLFSLMOYSEBFRIHFTFKFONNPNRSLVKPTQLLGRTHATGRKVVRELFN	240
CC		DB	181	KKSKTLFSLMOYSEBFRIHFTFNDKFNPSRSHVSPKQLNGRTKASGIRKVVRELFH	240
CC		QY	241	ITNGARKNAFKILVITDGEYFGDPLGYEDVIPADREGVIRYVIGVDADRSEKREL	300
CC		DB	241	KTNGARENAKILVITDGEYFGDPLGYEDVIPADREGVIRYVIGVDADRSEKREL	300
CC		QY	301	NTIASKPPRDHVFQVNNFEALKTIONOLREKI FAIEGTQTCSSSSSEHEMSEBGFSAIT	360
CC		DB	301	DTIASKPAGEHVFQVDNFEALNTIQNLOEKI FAIEGTQTCSTSSFEHEMSEBGFSAIT	360
CC		QY	361	SNGPLLSVGSYDAGGVFLYTSKEKSTFFINMTRVDSMDNDAYLGYAAAILRNVRQSLV	420
CC		DB	361	SNGPLLSVGSYDAGGVFLYTSKDKVTFINMTRVDSMDNDAYLGYASAVILRNVRQSLV	420
CC		QY	421	LGAPRYOHIGLVAMFRONTGMWESNANVKGTCIGAYFGASLCSVDVDSNGSTDLVIGAP	480
CC		DB	421	LGAPRYOHIGLVAMFRONTGMWESNANVKGTCIGAYFGASLCSVDMDADGNINLIGAP	480
CC		QY	481	HYEYTRGGQVSVCPPLPRGQARWQCDVAVLGEQGPWGRGAALTVLGDVNGDKLTDVA	540
CC		DB	481	HYEYTRGGQVSVCPPLPRG-BARWQCEALLHGDQHPWGRGAALTVLGDVNGDKLTDVA	539
CC		QY	541	IGAPGEDNRGAVLPHGTSGSGISPSHSQRIASKLSPLQYFGOSLSCGQDLTMDGLV	600
CC		DB	540	IGAPGEDNRGAVLPHGTSGSGISPSHSQRIASKLSPLQYFGOSLSCGQDLTMDGLM	599
CC		QY	601	DLTVGAQGHVLLRSQPVLRVKAIMEFNPREVARNVPCNDQVVKGEAGVRVCLHVQK	660
CC		DB	600	DLAVGAQGHVLLRSQPVLRVLEATWEPSPKVARSVFACQVQLKNDXGAGEVRVCLVRK	659
CC		QY	661	STRDLREGQISVYVTDALDSCRPHSRAVNETQNSSTRQTVLGLTQTCETIKLOLP	720
CC		DB	660	NTKDLREGQISVYVTDALDLPVRSIRAFEDTKNNTRRTQVFGLMQKCTIKLILP	719
CC		QY	721	NCIEEDVPSPVLRNLNFSIAGTPLSAFGNLRPVLAEDAQRLFTALPPPEKNCNDNICODD	780
CC		DB	720	DCVDSVSPVLRNLNFSIAGTPLSAFGNLRPVLAEDAQRLFTALPPPEKNCNDNICODD	779
CC		QY	781	LSITFSMSLDCLVGGPRFZNVTVVNDGEDSYRTQVTFPFLDLSYKRVKSLQNGRS	840
CC		DB	780	LSITFSMSLDCLVGGPRFZNVTVVNDGEDSYRTQVTFPFLDLSYKRVKSLQNGRPLT	839
CC		QY	841	QBSWPL-ACESASSTENSGALKSTSCSINHPIPEPENSEVTNITFDVDSKASLGNKLLK	899
CC		DB	840	KKPFVVKPAESSSSSEGGALKSTTWNINHPIFANSEVTNITFDVDSHASFNGKLLK	899
CC		QY	900	ANVTSENMPRTNTEFQLELPVKYAVVWVTVSHGVSTKYLNFTASENVSIRVMHQYQS	959
CC		DB	900	AIIVASENNMSRTHKTKFQLELPVKYAVVWVTVSHGVSTKYLNFTASENVSIRVMHQYQN	959
CC		QY	960	NLQBSLPISLVPLVPLVLAQTVIWDROPVTFPSENLSASTCHTKRELPHSHSFLAELRAP	1019
CC		DB	960	NLQBSLPISLVPLVPLVLAQTVIWDROPVTFPSENLSASTCHTKRELPHSHSFLAELRTP	1019
CC		QY	1020	VVNCISAVCQRIQCDIPFFGIQEEFNATLKNLSFDWYIKTSHNELIVSTAEILFNDVS	1079
CC		DB	1020	VVNCISAVCQRIQCDIPFFGIQEEFNATLKNLSFDWYIKTSHNELIVSTAEILFNDVA	1079

Query Match 76.3%; Score 4545; DB 1; Length 1153;
Best Local Similarity 74.5%; Pred. No. 9.5e-295;
Matches 860; Conservative 141; Mismatches 151; Indels 2; Gaps 2;

Query Match		99.7%; Score 5940.5; DB 1; Length 1152;
Best Local Similarity		99.9%; Pred. No. 0;
Matches 1152; Conservative 0; Mismatches 0; Indels 1; Gaps 1;		
QY	1	MALRVLLTALTLCGHNLDENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIIVAAQR 60
DB	1	MALRVLLTALTLCGHNLDENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIIVAAQR 60
QY	61	GSLYQCDYSTGCEPIRLQVPEAVNMSLGLSLAATTSPQLLACGPTVHTQCSNTYVK 120
DB	61	GSLYQCDYSTGCEPIRLQVPEAVNMSLGLSLAATTSPQLLACGPTVHTQCSNTYVK 120
QY	121	GLCFPGSNLRQOPKPFPEALRGCPQEDSDIAFLIDSGSIIPHPFRWKRPVSTVMEQL 180
DB	121	GLCFPGSNLRQOPKPFPEALRGCPQEDSDIAFLIDSGSIIPHPFRWKRPVSTVMEQL 180
QY	181	KSKTLPFLSMQSYSEPRFHTFKBPQNNPNSLVKPIITQLLGRTHATGIRKVVRELFN 240
DB	181	KSKTLPFLSMQSYSEPRFHTFKBPQNNPNSLVKPIITQLLGRTHATGIRKVVRELFN 240
QY	241	ITNGARKNAFKLLVITGCEKFGDPLGYEDVTPADREGVIRYVIGGDAFRSEKSRQEL 300
DB	241	ITNGARKNAFKLLVITGCEKFGDPLGYEDVTPADREGVIRYVIGGDAFRSEKSRQEL 300
QY	301	NTIASKPPRDHVFQVNNFEAKTIQNLREKIPALEGTQGTSSSSFEHMSQEGFSAIT 360
DB	301	NTIASKPPRDHVFQVNNFEAKTIQNLREKIPALEGTQGTSSSSFEHMSQEGFSAIT 360
QY	361	SNGLPSTVGSVDNAGGVFLYSKEKSTFINTRVDSMDNAYLYGAAAILLRNVQSLV 420
DB	361	SNGLPSTVGSVDNAGGVFLYSKEKSTFINTRVDSMDNAYLYGAAAILLRNVQSLV 420
QY	421	LGAPRYQHIGLVAMFRONTKWNESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLI GAP 480
DB	421	LGAPRYQHIGLVAMFRONTKWNESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLI GAP 480
QY	481	HYETRTGGQSVCPPLPGQARWOCDAVLIGEQQPMGRFGAALTVLGDVNGDKLT DVA 540
DB	481	HYETRTGGQSVCPPLPGQARWOCDAVLIGEQQPMGRFGAALTVLGDVNGDKLT DVA 539
QY	541	IGAPEENNRGAVLYFHTGSSGSPSHSRIAGSKLSPRLQYFGQSLGGQDITMDGLV 600
DB	540	IGAPEENNRGAVLYFHTGSSGSPSHSRIAGSKLSPRLQYFGQSLGGQDITMDGLV 599
QY	601	DLTVGAQGHVLLRSQPLRVKAIMFNPVRVARNVFECDQVWKGKEAGEVRVCLHVQK 660
DB	600	DLTVGAQGHVLLRSQPLRVKAIMFNPVRVARNVFECDQVWKGKEAGEVRVCLHVQK 659
QY	661	STRDLREGQIOSVVTYDLALDSGRPHSRVAFNETKSTRTQVGLGTQTCETLKLQLP 720
DB	660	STRDLREGQIOSVVTYDLALDSGRPHSRVAFNETKSTRTQVGLGTQTCETLKLQLP 719
QY	721	NCIEDPVSPIVLRNFSVLGTPLSAFGNLRPLVLAEDAQRFLTALFPFKKNCNDNI CDD 780
DB	720	NCIEDPVSPIVLRNFSVLGTPLSAFGNLRPLVLAEDAQRFLTALFPFKKNCNDNI CDD 779
QY	781	LSITFSFMSLCLVGGPRFNVTVTVNDGEDSVRTQVTFPFLDLSYRKVSTLQORS 840
DB	780	LSITFSFMSLCLVGGPRFNVTVTVNDGEDSVRTQVTFPFLDLSYRKVSTLQORS 839
QY	841	QRWRLACESASSTEVSGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKA 900
DB	840	QRWRLACESASSTEVSGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKA 899
QY	901	NVTSENMPRTNKTFFQLELPVKVAVVMTSHGVSTKYNLFTASENTSRVMOHQYQVSN 960
DB	900	NVTSENMPRTNKTFFQLELPVKVAVVMTSHGVSTKYNLFTASENTSRVMOHQYQVSN 959
QY	961	LQBSLPLSLVFLVPLVRLNQVTDQVTESENI SSTCHTKERLPHSDFLAELRKPV 1020
DB	960	LQBSLPLSLVFLVPLVRLNQVTDQVTESENI SSTCHTKERLPHSDFLAELRKPV 1019

QY	1021	VNC5IACVQRTQCDIPFGIOEBFNATLKGNSLSPDWIKTSHNHLIIVSTABILENDSVF 1080
DB	1020	VNC5IACVQRTQCDIPFGIOEBFNATLKGNSLSPDWIKTSHNHLIIVSTABILENDSVF 1079
QY	1081	TLLFGQAGFVSQETKVEPEVENPFLIVGSSVGGLLALITAAALYKLGFFKRYK 1140
DB	1080	TLLFGQAGFVSQETKVEPEVENPFLIVGSSVGGLLALITAAALYKLGFFKRYK 1139
QY	1141	MMSEGGPPGAPQ 1153
DB	1140	MMSEGGPPGAPQ 1152

RESULT 2

ID	ITAM MOUSE	STANDARD;	PRT; 1153 AA.
AC	POS555;		
DT	01-NOV-1988 (Rel. 09, Created)		
DT	01-FEB-1991 (Rel. 17, Last sequence update)		
DT	28-FEB-2003 (Rel. 41, Last annotation update)		
DE	Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor MO1).		
GN	ITGAM.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=88312584; PubMed=3044779;		
RA	Pytela R.;		
RT	"Amino acid sequence of the murine Mac-1 alpha chain reveals homology with the integrin family and an additional domain related to von Willebrand factor.";		
EL	EX30 J. 7:1371-1378(1988).		
RN	[2]		
RP	SEQUENCE OF 11-45 FROM N.A.		
RC	STRAIN=BALB/c; TISSUE=Spleen;		
RX	MEDLINE=86287312; PubMed=2942940;		
RA	Sastre L., Roman J.M., Teplow D.B., Dreyer W.J., Gee C.B.,		
RA	Larson R.S., Roberts T.M., Springer T.A.;		
RT	"A partial genomic DNA clone for the alpha subunit of the mouse complement receptor type 3 and cellular adhesion molecule Mac-1.";		
RL	Proc. Natl. Acad. Sci. U.S.A. 83:5644-5648(1986).		
RN	[3]		
RP	SEQUENCE OF 17-28.		
RX	MEDLINE=85198276; PubMed=3887182;		
RA	Springer T.A., Teplow D.B., Dreyer W.J.;		
RT	"Sequence homology of the LFA-1 and Mac-1 leukocyte adhesion glycoproteins and unexpected relation to leukocyte interferon.";		
RL	Nature 314:540-542(1985).		
CC	-1- FUNCTION: INTEGRIN ALPHA-M/BETA-2 IS IMPLICATED IN VARIOUS ADHESIVE INTERACTIONS OF MONOCYTES, MACROPHAGES AND GRANULOCYTES AS WELL AS IN MEDIATING THE UPTAKE OF COMPLEMENT-COATED PARTICLES. IT IS IDENTICAL WITH CR-3. THE RECEPTOR FOR THE IC3B FRAGMENT OF THE THIRD COMPLEMENT COMPONENT. IT PROBABLY RECOGNIZES THE R-G-D PEPTIDE IN C3B. INTEGRIN ALPHA-M/BETA-2 IS ALSO A RECEPTOR FOR FIBRINOGEN, FACTOR X AND ICAM1. IT RECOGNIZES P1 AND P2 PEPTIDES OF FIBRINOGEN GAMMA CHAIN. ALPHA-M/BETA-2 PLAY A CRITICAL ROLE IN MAST CELL DEVELOPMENT AND IN IMMUNE COMPLEX-MEDIATED GLOMERULONEPHRITIS. MICE EXPRESSING A NULL MUTATION OF THE ALPHA-M SUBUNIT GENE DEMONSTRATE INCREASE IN NEUTROPHIL ACCUMULATION, IN RESPONSE TO A IMPAIRED DEGRANULATION AND PHAGOCYTOSIS, EVENTS THAT APPARENTLY ACCELERATE APOPTOSIS IN NEUTROPHILS. THESE MICE DEVELOP OBESITY.		
CC	-1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN. ALPHA-M ASSOCIATES WITH BETA-2.		
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein.		
CC	-1- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND GRANULOCYTES.		
CC	-1- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.		
CC	-1- SIMILARITY: Belongs to the integrin alpha chain family.		

RP SEQUENCE OF 1-9 FROM N.A.
RX MEDLINE=92073318; PubMed=1663702;
RA Shelley C.S., Arnaut M.A.;
RT "The promoter of the CD11b gene directs myeloid-specific and
RT developmentally regulated expression.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:10525-10529(1991).
RN [7]
RP SEQUENCE OF 1-9 FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=92144986; PubMed=1346576;
RA Paul H.L., Rosmarin A.G., Tenen D.G.;
RT "Characterization of the myeloid-specific CD11b promoter.";
RL Blood 79:865-870(1992).
RN [8]
RP SEQUENCE OF 17-31.
RX MEDLINE=87076671; PubMed=3539202;
RA Pierce M.W., Remold-O'Donnell E., Todd R.P. III, Arnaut M.A.;
RT "N-terminal sequence of human leukocyte glycoprotein Mol.
RT conservation across species and homology to platelet IIb/IIIa.";
RL Biochim. Biophys. Acta 874:368-371(1986).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 148-331.
RX MEDLINE=95171458; PubMed=7867070;
RA Lee J.O., Rieu P., Arnaut M.A., Liddington R.;
RT "Crystal structure of the A domain from the alpha subunit of integrin
RT CR3 (CD11b/CD18).";
RL Cell 80:631-638(1995).
RN [10]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 148-334.
RX MEDLINE=96363671; PubMed=8747460;
RA Lee J.O., Bankston L.A., Arnaut M.A., Liddington R.C.;
RT "Two conformations of the integrin A-domain (I-domain): a pathway for
RT activation?";
RL Structure 3:1333-1340(1995).
RN [11]
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 148-337.
RX MEDLINE=98362595; PubMed=9687375;
RA Baldwin E.T., Sarver R.W., Bryant G.L. Jr., Curry K.A.,
RA Fairbanks M.B., Finzel B.C., Garlick R.L., Heinrikson R.L.,
RA Horton N.C., Kelley L.L., Milner A.M., Moon J.B., Mott J.E.,
RA Muchler V.T., Tomich C.S., Watenpaugh K.D., Wiley V.H.;
RT "Cation binding to the integrin CD11b I domain and activation model
RT assessment.";
RL Structure 6:923-935(1998).
RN [12]
RP 3D-STRUCTURE MODELING OF 17-616.
RX MEDLINE=98226734; PubMed=9560195;
RA Oxvig C., Springer T.A.;
RT "Experimental support for a beta-propeller domain in integrin alpha-
RT subunits and a calcium binding site on its lower surface.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:4870-4875(1998).
CC -!- FUNCTION: INTEGRIN ALPHA-M/BETA-2 IS IMPLICATED IN VARIOUS
CC ADHESIVE INTERACTIONS OF MONOCYTES, MACROPHAGES AND GRANULOCYTES
CC AS WELL AS IN MEDIATING THE UPTAKE OF COMPLEMENT-COATED PARTICLES.
CC IT IS IDENTICAL WITH CR-3, THE RECEPTOR FOR THE IC3B FRAGMENT OF
CC THE THIRD COMPLEMENT COMPONENT. IT PROBABLY RECOGNIZES THE R-G-D
CC PEPTIDE IN C3B. INTEGRIN ALPHA-M/BETA-2 IS ALSO A RECEPTOR FOR
CC FIBRINOGEN, FACTOR X AND ICAM1. IT RECOGNIZES P1 AND P2 PEPTIDES
CC OF FIBRINOGEN GAMMA CHAIN.
CC -!- SUBUNIT: HETERO DIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-M
CC ASSOCIATES WITH BETA-2.
CC -!- SUBCELLULAR LOCATION: Type 1 membrane protein.
CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND
CC GRANULOCYTES.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VMPA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC -!- SIMILARITY: Contains 1 VMPA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD11b entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd11b.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J03925; AA59544.1; -
DR EMBL; M18044; AA59491.1; -
DR EMBL; J04145; AA59903.1; -
DR EMBL; S5227; AAB34821.1; -
DR EMBL; S52152; AAB34821.1; JOINED.
DR EMBL; S52153; AAB34821.1; JOINED.
DR EMBL; S52154; AAB34821.1; JOINED.
DR EMBL; S52155; AAB34821.1; JOINED.
DR EMBL; S52157; AAB34821.1; JOINED.
DR EMBL; S52159; AAB34821.1; JOINED.
DR EMBL; S52161; AAB34821.1; JOINED.
DR EMBL; S52164; AAB34821.1; JOINED.
DR EMBL; S52165; AAB34821.1; JOINED.
DR EMBL; S52167; AAB34821.1; JOINED.
DR EMBL; S52169; AAB34821.1; JOINED.
DR EMBL; S52170; AAB34821.1; JOINED.
DR EMBL; S52173; AAB34821.1; JOINED.
DR EMBL; S52174; AAB34821.1; JOINED.
DR EMBL; S52180; AAB34821.1; JOINED.
DR EMBL; S52181; AAB34821.1; JOINED.
DR EMBL; S52184; AAB34821.1; JOINED.
DR EMBL; S52185; AAB34821.1; JOINED.
DR EMBL; S52191; AAB34821.1; JOINED.
DR EMBL; S52192; AAB34821.1; JOINED.
DR EMBL; S52203; AAB34821.1; JOINED.
DR EMBL; S52212; AAB34821.1; JOINED.
DR EMBL; S52213; AAB34821.1; JOINED.
DR EMBL; S52215; AAB34821.1; JOINED.
DR EMBL; S52219; AAB34821.1; JOINED.
DR EMBL; S52220; AAB34821.1; JOINED.
DR EMBL; S52221; AAB34821.1; JOINED.
DR EMBL; S52222; AAB34821.1; JOINED.
DR EMBL; S52226; AAB34821.1; JOINED.
DR EMBL; M76724; AA58410.1; -
DR EMBL; M84477; AA51960.1; -
DR PIR; A31108; RHUUB.
DR PDB; 1ABX; 17-JUN-98.
DR PDB; 1BHO; 18-NOV-98.
DR PDB; 1BHQ; 18-NOV-98.
DR PDB; 1IDN; 25-NOV-98.
DR PDB; 1IDO; 01-AUG-96.
DR PDB; 1JLM; 11-JAN-97.
DR PDB; 1MIU; 07-AUG-02.
DR Genew; HGNC:6149; ITGAM.
DR MIM; 120980; -
DR GO; GO:0008305; C:integrin complex; TAS.
DR GO; GO:0004895; F:cell adhesion receptor activity; TAS.
DR GO; GO:0007155; P:cell adhesion; TAS.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWFA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Signal; 3D-structure; Repeat; Magnesium; Calcium.
FT SIGNAL 1 16
FT CHAIN 17 1152 INTEGRIN ALPHA-M
FT DOMAIN 17 1104 EXTRACELLULAR (POTENTIAL)
FT TRANSMEM 1105 1128 POTENTIAL
FT DOMAIN 1129 1152 CYTOPLASMIC (POTENTIAL)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 16:55:38 ; Search time 10.1123 Seconds
(without alignments)
5937.039 Million cell updates/sec

Title: US-09-902-481b-1
Perfect score: 5956
Sequence: 1 MALRVLLTALTLCGHNLD.....FKROYKDMSEGPGCAEPQ 1153

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5940.5	99.7	1152	1 ITAM_HUMAN	P11215 homo sapien
2	4540.5	99.7	1152	1 ITAM_MOUSE	P05555 mus musculus
3	3503	58.8	1163	1 ITAD_HUMAN	P20702 homo sapien
4	3455	58.0	1162	1 ITAD_HUMAN	Q13349 homo sapien
5	1563	26.2	1170	1 ITAL_HUMAN	P20701 homo sapien
6	1549	26.0	1163	1 ITAL_MOUSE	P24063 mus musculus
7	1165.5	19.6	1167	1 ITAE_MOUSE	Q60677 mus musculus
8	1163	19.5	1179	1 ITAE_HUMAN	P38570 homo sapien
9	1108.5	18.6	1151	1 ITAL_HUMAN	P56199 homo sapien
10	1102.5	18.5	1189	1 ITAH_HUMAN	Q9UKX5 homo sapien
11	1099	18.5	1170	1 ITA2_BOVIN	P53710 bos taurus
12	1089	18.3	1178	1 ITA2_MOUSE	Q62469 mus musculus
13	1087.5	18.3	1181	1 ITA2_HUMAN	P17301 homo sapien
14	1085.5	18.2	1167	1 ITAG_HUMAN	Q75578 homo sapien
15	1082	18.2	1180	1 ITAL_RAT	P18614 rattus norv
16	667	11.2	1039	1 ITA4_MOUSE	Q00651 mus musculus
17	638	10.7	1038	1 ITA4_HUMAN	P13612 homo sapien
18	630	10.6	1035	1 ITA9_HUMAN	Q13797 homo sapien
19	600	10.1	1032	1 ITA4_XENLA	Q91687 xenopus lae
20	571.5	9.6	1066	1 ITA3_CRISP	P17852 cricetidae
21	567.5	9.5	1053	1 ITA3_MOUSE	Q62470 mus musculus
22	555.5	9.3	1053	1 ITA5_MOUSE	P11688 mus musculus
23	550	9.2	1034	1 ITAV_CHICK	P26008 gallus gall
24	546.5	9.2	1050	1 ITA5_XENLA	Q06274 xenopus lae
25	540	9.1	1130	1 ITA6_HUMAN	P23229 homo sapien
26	535	9.0	1044	1 ITAV_MOUSE	P43406 mus musculus
27	532	8.9	1049	1 ITA5_HUMAN	P08648 homo sapien
28	531.5	8.9	1066	1 ITA3_HUMAN	P26006 homo sapien
29	530.5	8.9	1072	1 ITA6_CHICK	P26007 gallus gall
30	526	8.8	1048	1 ITAV_HUMAN	P06756 homo sapien
31	517	8.7	1044	1 ITA8_CHICK	P26009 gallus gall
32	512.5	8.6	1091	1 ITA6_MOUSE	Q61739 mus musculus
33	508	8.5	1396	1 ITA2_DROME	P12080 drosophila

RESULT 1				
ITAM_HUMAN	STANDARD;	PRT;	1152 AA.	
ID	ITAM_HUMAN			
AC	P11215;			
DT	01-JUL-1989 (Rel. 11, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor MOI)			
DE	(Neutrophil adherence receptor).			
GN	ITGAM OR CR3A OR CD11B.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88315033; PubMed=2457584;			
RA	Corbi A.L., Kishimoto T.K., Miller L.J., Springer T.A.;			
RT	"The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CD11b) alpha subunit. Cloning, primary structure, and relation to the integrins, von Willebrand factor and factor B.";			
RL	J. Biol. Chem. 263:12403-12411(1988).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88190151; PubMed=2833753;			
RA	Arnaout M.A., Remold-O'Donnell E., Pierce M.W., Harris P., Tenen D.G.;			
RT	"Molecular cloning of the alpha subunit of human and guinea pig leukocyte adhesion glycoprotein Mol: chromosomal localization and homology to the alpha subunits of integrins.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 85:2776-2780(1988).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88257215; PubMed=2454931;			
RA	Arnaout M.A., Gupta S.K., Pierce M.W., Tenen D.G.;			
RT	"Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor Mol (complement receptor type 3).";			
RL	J. Cell Biol. 106:2153-2158(1988).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=93123748; PubMed=8419480;			
RA	Fleming J.C., Pahl H.L., Gonzalez D.A., Smith T.F., Tenen D.G.;			
RT	"Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-integrin gene family demonstrate remarkable conservation of genomic organization and suggest early diversification during evolution.";			
RL	J. Immunol. 150:480-490(1993).			
RN	[5]			
RP	SEQUENCE OF 9-1153 FROM N.A.			
RX	MEDLINE=89098993; PubMed=2563162;			
RA	Hickstein D.D., Hickey M.J., Ozols J., Baker D.M., Back A.L., Roth G.J.;			
RT	"cDNA sequence for the alpha M subunit of the human neutrophil adherence receptor indicates homology to integrin alpha subunits.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 86:257-261(1989).			
RN	[6]			

34	498	8.4	1179	1	ITA7_MOUSE	Q61738 mus musculus
35	492	8.3	1033	1	ITAB_MOUSE	Q9GUM0 mus musculus
36	491.5	8.3	1146	1	ITAI_DROME	Q24247 drosophila
37	489.5	8.2	1039	1	ITAB_HUMAN	P08514 homo sapien
38	489	8.2	1025	1	ITAB_HUMAN	P53708 homo sapien
39	486	8.2	126	1	ITAM_CAVPO	P11578 cavia porce
40	474.5	8.0	1226	1	PAT2_CAEEL	P34446 caenorhabdi
41	472	7.9	1181	1	ITAT_HUMAN	Q13683 homo sapien
42	469.5	7.9	1106	1	ITAT_RAT	Q63258 rattus norv
43	454.5	7.6	1139	1	INAI_CAEEL	Q03600 caenorhabdi
44	424.5	7.1	1115	1	ITAI_DROME	O44386 drosophila
45	383	6.4	1000	1	ITAS_DROME	Q9WIM8 drosophila

ALIGNMENTS

Db 702 DFL-----KCSVGFFPMRSKSYEFVDPDTHLS--GEEVLSPIVTAQSG--NTER 750
 Qy 914 TE-----FQLELPVKYAVVWVTS-----HGVSTKYLNFTASNTS---RVMQHOY 956
 Db 751 SESLHDVTLVLMFMEHVDTSITGIMSPISFVIGESVDAAFIQLDLECHFPQINIL 810
 Qy 957 QVSNLQSRSLPISLVL--VPVRLN-----QTVWDRPQVPTFSENLSSTCHTKER 1004
 Db 811 QVANTGPSTLPGSSVSISFNNRLSSGGAEMFHVQEMVVGQKGNCSFQKNPTPCIIPOEQ 870
 Qy 1005 LPSHSDFLAELRKA-----PWNCSIAVCQRIQCDIPFFGIGQEFNATLKGNSPD 1055
 Db 871 ENIFHTTIFATPKGRKVLDCKEGICSLTHCN-----FSALEBESRTI-----D 917
 Qy 1056 WYIKTSHNHLIVSTAEILFNDSTFTLLPGQCAFVRSQTERKVP----- 1100
 Db 918 IY-----MLLAT--ELKKDSSVIQ-----FMSRAKVKVDPALRVVEIAHGNPEEV 962
 Qy 1101 ---PEVPN---PLPLIVG-----SSVGLLLALITAILYKLGFPKQYKMM 1142
 Db 963 TVWFEALHNLPRGVVGVWIIAISLLGILIFLLVLLVMKMGFRRRYKEII 1015

RESULT 15
 T31437
 integrin alpha chain SU2 - sea urchin (Lytechinus variegatus)
 C:Species: Lytechinus variegatus (variegated urchin)
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jan-2000
 C:Accession: T31437
 R:Hertxler, P.L.; McClay, D.R.
 submitted to the EMBL Data Library, May 1998
 A:Description: Alpha SU2, a sea urchin integrin which binds laminin.
 A:Reference number: Z21035
 A:Accession: T31437
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-1041 <HER>
 A:Cross-references: EMBL:AF067658; NID:G3220240; PID:G3220241; PIDN:AAC23572.1
 A:Experimental source: developmental stage embryo
 C:Function:
 A:Description: binds laminin
 C:Superfamily: integrin alpha-2b chain

Query Match 10.3%; Score 614.5; DB 2; Length 1041;
 Best Local Similarity 24.8%; Pred. No. 4e-34;
 Matches 236; Conservative 155; Mismatches 314; Indels 245; Gaps 41;

Qy 352 QEGFSAITNGP--LLSTVGSYDWAGGVFLYTSKESKSTFINMTRVDS-----DMNDAYLG 405
 Db 181 QAGPSGIIFSDNSALVNGAFSGSYLQCIYVQSLLNRSV--VQATQESNTGTYSFDNSYRG 239
 Qy 406 YAAAI--ILNRVQSLVLAGAPRYOHI--GLVAMPFRONTGMWESNANVKGTQIGAYFGASLC 462
 Db 240 YSLALGPFNGDGVQDYVGTGFAESLMLGLVAIFQNLNQFN---QVMSTQIVAYFGYSVT 296
 Qy 463 SVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPPLRQORARWQCDVLYGEQ----- 514
 Db 297 VVDI--NNDTYDILLVAGAPMYMDGPAIQ-----RWEAGAVVYVQLQNPVGPGA 343
 Qy 515 -----GQWGRFGAALTVLVDVNGDKLTDVAICARGEEDNRGAVLPHGTSGS 562
 Db 344 SNLSLSSTLIGQIRERFGLSTASIGDSNQDGFNDVAIGAPYEGDAGAVIYHG--SAN 402
 Qy 563 GISPSHSQRIAGSKLS--PRIYFGQSLSGGQDLTMDGLVLTGVAQ--GHVLLRSQPVL 619
 Db 403 GLKSTPAQVLTPTSLGHSGITTFGFSLQGGQDMKKNYPDLLVGAESANTAVLIRTPVV 462
 Qy 620 RVKALMEFNPREVARNVFCNDQVWVKEAGEVRVCLHVQKSTEDRLREGQIOSV----- 674
 Db 463 SLDATLNTPE-----IGINLENKTYB--LADGTMTVTSFIAMT 497
 Qy 675 -----VTYDLALDSG--RPHSRAVFNETKNSSTRQTVIGL--TOTCETILKLQ 718

Search completed: June 7, 2004, 17:17:33
 Job time : 20.7639 secs

Db 498 CFTYTGNYLPDHIDISVTVTVDSGIIANRRAMFVNDMDSEITKTRRLAVSTQPCDPLRAY 557
 Qy 719 LPNCIEDPVSPIVLRNLNFSIVCTPLSAGFN-----LRPVLAEDAQRLFTALFPPEK 769
 Db 558 VGNISIEDKUTPITKVTIYQIDL-----NNDSERLQPHLPIIDMATSTQTKQVSIQN 609
 Qy 770 NCNDNICODDLISITFSFMSLDCLVVGGPREFNVTVVRNDGSDSYRTQVTFPPDLDSY 829
 Db 610 NCVN--NICIPDLVDVTV--PNLPNVIGQTOELTLDVSLANRGEDAFQSSLSVYVPLGLQF 667
 Qy 830 RKYSTLQNRQSRWRLACESASSTEVSGALKSTCSINHPIPPEN-----SEVFNIT 893
 Db 668 VRL-----ERKANMDFSVTCESD-----LRITCDTGNPMVGNKILBFGLTSTFQVS 717
 Qy 884 FDVDS-----KASLGNKLLKLLKANVTSENMMPTINKTEFOLELPVKYAVVMTSHGVSTK 938
 Db 718 GDKDSIEFYFKASENS--EDPNTLENELNMTVPVTVDTLKLKLSASYPEIWMYSTQED 775
 Qy 939 YL-----NFTASENTSRVMQHOYQV-----SNLQORSILPIS-----LVFLVPV 976
 Db 776 YVVPFPKAKNASEADIGMEVM--HLXEVRNTGSGNAGEVSLNIQWPOKNEDEGYLFYLLGI 834
 Qy 977 RLNQTVIWDRPQ-----VTFSENLSSTCHTKERLPSESDPLAELRKAPVV 1021
 Db 835 MTEEGVTCOLTOGKANPEGVKLEPSTKAKLSNSTTQVSGRKRPEPEVAEALQTDN--VI 992
 Qy 1022 NCSIAVCQRIQCDIPFFGIGQEFNAT-----LKGNLSTFDWYIKTSHNHLIVSTAEILF 1075
 Db 893 YCASDSCVLINCTI-----DEINASKSKVAILGRF--W-----ERTF 928
 Qy 1076 NDSVFLLPQOGAFVRSQTERKVE--PPEVPNP-----LP----- 1108
 Db 929 QKAVSELTPWQVQTIASSASAAVKTIPIPLPRDPSDSTKASTLVTTEELVPPVTPIAW 998
 Qy 1109 -LIVSGSVGLLLALITAILYKLGFPKR-----QYKDMMS--EGGPP 1148
 Db 989 WIIVSVLGGIILMLIIILGLWKCGFFERKPKGEKEVAPVASADKGGP 1038

Qy	178	EOLX--KSKYLFSLMOYSEFRHFTFKFQNNPNPRLVKPITOLL----	GRTHATGI	231
Db	189	QGLDIGETKTQMGILQIYANNPRVFNLTFKSKD--EMIKATSQTQYQGD	LNTFKAI	245
Qy	232	RKVRELFNITNGARKNAFKILVITDGEKFGDPLGYEDVTPBADRGVIRYIGV----	287	
Db	246	QYARDTAYSTAAAGRPGATKVVVVVTDGESH--DGSKLKAVIDCOKNIDRPGVLAVLYL	304	
Qy	288	GDAPFRSEKROBLNTTASRPRDHVQVNNFBAKTIQNLREKFIPIAIGTOTGSSSSP	346	
Db	305	NRNALDPTNLIKETIKATASIPTEHRFPFNVSDEADLLEKAGTIGEQIFSIETGVQ--	GDNF	363
Qy	347	EHMSQBGFAAIT--SNGPLLSVTGSDWAGGVFLYTSKEKSTFINMT--RVDSDKN-D	401	
Db	364	QMEKSQVGFABYSPPNNILMLGAVGAYDWSGVVQKTPHGHLFSKAQEILQDRNHS	423	
Qy	402	AYLGAAAAIILNRVQSLVIGAPRYQHILGVMPRQNTGMWESNANV-----	KGTGIGA	455
Db	424	SYLGYSVASISTGNSVHFAGAPRANTGTQIIVSYVN-----ENGNTVVTQSORQDQIGS	478	
Qy	456	YFGASICSVDVDSNGSTDVLIGAPHYYEQTR--GGOVSVCPILPRQARARWQCDVILYE	513	
Db	479	YFGSVLCAVDVNDKTTIDVLLVGAPMTMDLKKEEGRYVLTITKG--ILNWH--QFLEGP	535	
Qy	514	QOQPMWGRGAALTIVLGVNDGKLTDAIGAPBEDNRGAVYLFHGTSGSIGSPSHSORIA	573	
Db	536	NGLENARFGSAIAALSIDINDGFNDVTGVSPLENQNSGAVIYNGEHGM--IRLRYSKIL	594	
Qy	574	GS--XLSPLRIYFGOISGGODLTMDGLVBLTVGAQGHVILLLRSPQVRLVKAIWEFNPRE	631	
Db	595	GSDRAPSSHLOYFGRSLDYGDLNGDSITDVSUGAFQVQVLWQSQTADVSVNDASFTFKK	654	
Qy	632	VARNVFCNDQVVKGEAGVVRVCLHVQKSTRDKLREGQIQSVVYTDLALD----	SGRPH	687
Db	655	I--TLLNKNAEI-----KLKLCF-----SAKFRPTQNNQVAIVYITIDEDQFSSRVI	701	
Qy	688	SRAVFNETKNSRROTQVLGLTQICE--TLKLQLPNCIEDVPSPVILRLNFSI--VCTPL	743	
Db	702	SRGLFKENNERCLKNTWIVSQACRSYIITHIQEPS--DIISPLNOMISLENPGT--	756	
Qy	744	SAPGNLRPLVLAEDAQRLFTALFFPERKCGNDNICQDDLSITF-----SFMSLDCLVVGGR	799	
Db	757	-----NPALFAYSETVKVPSIPHHKCGDGDGVCISDLVLNVQQLPATQOQPFIVSNQNK	810	
Qy	800	EFNVTVTRVNDGDSYETQVTFPPFPLDLSYRKVSTLQNRQSORSWRLACESAST--EVSQ	858	
Db	811	RLTFSVLQKNGKESAYNTEIIVVDSENLFF-----ASWSMPVDGTEVTCQIAS	858	
Qy	859	ALKSTSCSINHPIEPENSEVTFNITPFDVDSKASIGNKLLKANVTSENMPRTNKTBFOL	918	
Db	859	SQKSVTCNVGVPALKSKQVTFITNPFNLQ--NLONQASISFRALSSQENMADNSVNL	917	
Qy	919	ELPKVYAVMVVTSHGVSCTKLNFNTASENTRVMOHQYQVSNLQOR-----SLPI	968	
Db	918	KLSLYDARHIT--RSTNINPFYSLDGNVSSV--HSFE--DIGPKTFISKVITGSPVP	973	
Qy	969	SLVFLVPLVRLMQTVIWRDPQVTFSEN--LSSTCHTKE-----RLPSHS	1009	
Db	974	SMA-----SVIHHIPLYTKDKNPLMYLTGVHTDQAGDISCEASINPLKIQOTSSV	1024	
Qy	1010	DFLAE--LRKAPVNCSTAVCORICQDIPFGIQEENFATLKNLSPDWYIKTSHNLLIV	1068	
Db	1025	SFXSENFRHIKELNCRATASCNSINCMWLRLDQVKGGEYFLNSTRWNCTFAASTFQTVOLT	1084	
Qy	1069	STABE--LPNDSVFTL-----LPQCGAFVRSOTETKVPPE--VBNPLVLVSGSVGLLL	1120	
Db	1085	AAABIDYNQIYVIBENTVTP-----LTFMKPHEKVEVTPGVIVGSVIAGILL	1134	
Qy	1121	LALITALYKLGFFPKQYKDM	1141	
Db	1135	LLALVAILWKLGFFPKRYKYM	1155	

RESULT 9

S44142
VIA-2 protein homolog - mouse
C:Species: Mus musculus (house mouse)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 15-Sep-2003
R:Edelman, J.M.; Chan, B.M.; Uniyal, S.; Onodera, H.; Wang, D.Z.; Damjanovich, L.; Lat;
submitted to the EMBL Data Library, January 1994
A:Description: The mouse VIA-2 homologue supports collagen and laminin adhesion but not
A:Reference number: S44142
A:Accession: S44142
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1178 >E>
A:Cross-references: EMBL: Z29987; MID: g473098; PIDN: CAA82877.1; PID: g473099
P:169-344/Domain: von Willebrand factor type A repeat homology <WAA>

Query Match 18.3%; Score 1089; DB 2; Length 1178;
Best Local Similarity 28.1%; Pred. No. 4.8e-67;
Matches 348; Conservative 209; Mismatches 492; Indels 188; Gaps 44;

Qy	5	VLLALTLALCHGNLDTENAWTQ--ENARFGSGSVQL---QGSRVVVGAPQEIIVANQR	60
Db	15	LMVLQGIINCLAVNVLPGAKISGSPSEQFGYSVQOLTPTQGNLLWLVGSPGSPFPERM	74
Qy	61	GSLYQC--DYSTGSCBPIRLQ-----VFPEAVNMSLGLSLAATTSPPOLLACGPTVHQ	111
Db	75	GDVVKCPVDLPTATCEKLNLSNGASISNVTEIKTNMSLGLTLTRNPGTGGLTCGPLWAH	134
Qy	112	TCSENTYVKGCLFGSNLRQQPQ---KPEALRGCPQEDSDIAFLIDGSGSIIPHDFRR	168
Db	135	OCGNQYVATGIC---SDVSDFQPLTSPGPAQACPSL--VDVVVWCDENSIYP--WEA	187
Qy	169	MKEPFSVTWMEQLK--KSKTLFSLMQSSEBPIRIHFTFKBFQNNPNRSLVXPIITLLG--RT	225
Db	188	VKNFLVKFVTGLDIGPKKTKVALIQVANEPIRIIFNLNDFETKEDMWQATSETRCHGGDLT	247
Qy	226	HTATGIRKVRBELFNITNGARKAPKLLVITGCKFGDPLGYEDVTPADREGVIRVI	285
Db	248	NTRAFEFARDYAYSQTSGRGPGATKVMVVVTGESH--DGSKLTVIQCCNDDEILRPGI	306
Qy	286	GV-----GDAPFEKSRQELNTASKPPRDHVQVNNFEALKITIQNLREKIFAIEGTQT	340
Db	307	AVLGYLNRNALDKNLKIKAIKASTPTERYFENVADAEALKEKAGTLGEQIFSIGTVQ	366
Qy	341	GSSSSPHEMSQGFSA--AITNGPLLSTVGSYDRAGGVFLYTSKESKSFINMT--RVD	396
Db	367	G-GDNFQEMAWAQGFSADYAPNDIIMLGAVGAFDMSGTLLVQETSHKPEVFPKQAFQVL	425
Qy	397	SDMN--DAYLGAATAIILRNRSQSLVLCAPRYQHIGLVAMFRQNTGMWESNANV---KGT	451
Db	426	QDRNHSSFLGYSVAALSTEDGVHFVAGAPRANYTGQIVLYSVNK---QGNVTVIQSHRGD	482
Qy	452	QIGAYFGASICSDVDVDSNGSTDVLICAPHYIQTR--GGQVSVCPPLPRGARWQCDVAV	509
Db	483	QIGSYFGSVLCSVDVDKDTITDVLVAGAPTAMDLLKKEGKVLFTTKGILNQHFQ---F	539
Qy	510	LYEGOGPWRFGAALTVLGDVNGDKLTDVAIGAPGEEDNKGAVYLFHGTSGSGISPSHS	569
Db	540	LEGPEGTGNARFGSAIINALSDINWDGENDVIVKSPVENENSGAVIYINGHQT--IRKYS	598
Qy	570	QRIAGSKLSPR--LQYFGSLSGQDUTMDGLVDLTVAGACHVLLLSQPVLRYKAIWEP	627
Db	599	QKILGNSGAFRRHLQPFGRSLDGYGLDNGDSITDVSIGALCQVTLQMSQSIAADVAIBALF	658
Qy	628	NPREVARNVECDNDVVVKGEAGEVRVCLHVOKSTRDLREGQIQSVVTVYDLAID---S	683
Db	659	TP-----DKITLAKDKAKITLKCFAEFPFAGQNNQV---AILENMTLADGHS	705
Qy	684	GRHSRAVFNETKNSTRQTVLGLTQTCET--LKLQLPNCIEDPVPSPVILRLNFSLVGT	741
Db	706	SKVTSRGVFNENSRFLQKNVWVNEVKCEHHISIOKPS---DVNPLDLRVDISLENP	762

QY 1107 LPLIVGSSVGGHLLALITAAALYKLGFFKQYKDMSE 1144
 Db 1125 LPLIIGSVGGHLLVILVILFKGPFKRYOOLNLE 1162

RESULT 7
 A45226
 integrin alpha-1 chain - human (fragment)
 C:Species: Homo sapiens (man)
 C>Date: 30-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 15-Sep-2003
 C:Accession: A45226
 R:Briesewitz, R.; Epstein, M.R.; Marcantonio, E.E.
 J. Biol. Chem. 269, 2989-2996, 1993
 A:Title: Expression of native and truncated forms of the human integrin alpha 1 subunit.
 A:Reference number: A45226; MUID:9315124; PMID:8428973
 A:Accession: A45226
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-1151 <BRI>
 A:Experimental source: hepatoblastoma cell line HepG2
 A:Note: sequence extracted from NCBI backbone (NCBIP:124326)
 F:142-317/Domain: von Willebrand factor type A repeat homology <WAI>

Query Match 18.6%; Score 1108.5; DB 2; Length 1151;
 Best Local Similarity 28.0%; Pred. No. 2e-68;
 Matches 346; Conservative 210; Mismatches 486; Indels 195; Gaps 44;

QY 17 ENLDTENAMTFOENARG-FQGSVVOL---QGSRVVVGAPQEIIVANORGLYQCDYSTGS 72
 Db 1 FNVVKNMTFSGVEDMPGTYTQVYENBEGKWLGLSGPLVGQPKRTGVDYKCPVGRGE 60

QY 73 CEP-IRLOVPEA-----VNMSLGLSLAATTPPQLLACGPTVHOTCSENITVVKGL 122
 Db 61 SLPCVKLPLPNTSIPNVTVEKNMTFGSL-VTNENGFLACGLPYAYRQGHLYTGI 119

QY 123 CFLFGNLRQOPKPPALRGCPQEDSDIAFLDGGSGSIIPHFRMKPFVSTVMEQLK- 181
 Db 120 CSDVSPTFQVNSIAP--VOECSTQ-LDIVIVLDGNSIYPMD--SVTAFNLNLRMDI 174

QY 182 -KSKTLFSLMOYSEPRHFTKFEKONNPNRSLVPKIPITOLLGR-THTATGIRKVVRLP 239
 Db 175 GPKQTQGVIGYGENVTHEFNLYKSYSTEVLVAKKIVQGRGQRTWLTGDTARKEAF 234

QY 240 NITNGARKNAFLVITGKEGDPGLVEDVTPADRGVIRYIVTGVDAPR-----SE 294
 Db 235 TEARGARRGVKKVWVITDGESH-DNHRLLKVIQDCEDENIQRFSAIILGYSNRGNLSTE 293

QY 295 KSRQELNTIASKEPRDHPVQVNNFEALKTIONOLREKIFALEGTOVGSSSSSEHEMSQEG 354
 Db 294 KPVVEIKSIASEPTEKHPFNVSDELALVTIVKTLGRIFALEATADQSAASEFEMENSQTG 353

QY 355 FSAATISNGPLSTVGSYDWAGGVFLYTSKE-----KSTP-INMTVDSMDMDAVLYGAA 408
 Db 354 FSAHYSDQVWMLGAVGAYDWGTVVMQKASQIIIPRNTTFNVESTKKNEL-ASYLGIVT 412

QY 409 AILIREVOSL-VLGPAPYOHIGLVAMFRQNTGWESNANVKGTOIGATFGASLCSVDV 467
 Db 413 NSATASSGDVLVIAGOPRYNHTGQVIYRMEDGNIKILOTLSEGIQSYFGSILATTDID 472

QY 468 SNGSTDVLVLIGAPHY-----YEQTR-GGQVSVCPPLPRGQARWQCDQAVLY 511
 Db 473 KDSNTDILLVAGPMYWGTEKEGQKVVYVYALNQTRFYMKSLEPIKQTCSSRQHSCTT 532

QY 512 GQCGQWAG-RFGMALTVLDVNGDKLTVAGAPGEEDNRGAVYLFHGTSGSGISPSHSQ 570
 Db 533 ENKNEPCGARFETAIAAVKDLNLDGNDIVIGAPLDDHGGAVIYHG-SGRTIRKEVAQ 591

QY 571 RIAGSKSLPRLOVFGOSLGGODLTMDGLVDLVGAQGHVLLRSLQPLRVKAIEMENPR 630
 Db 592 RIFSGDGKTLFFGOSIINGEMDNGDLTVTIGLGAALFWERDVAIVKVTWTFNFBP 651
 QY 631 EVARNVFECDNQVYKGEAG--EVRVCLHVQ-KSTRDLRLBQIQSVVYDYLALDSGRPH 687

Db 652 KVNIOKKNCH---MEKETVCINATVCFVKLSKEDTITYEADLQ-----YRVTLDSLRQI 704
 QY 688 SRAVFNET-----KNSTRROTQVLGLTQTCRTKLQLPNCIEDPVSPIVLRNPLSLVGT 741
 Db 705 SRSPFSGTOERKVQRNITVRKSEC-----TKHSFYMLDKHDFQDSVR---IYLDFNLT-D 755

QY 742 ELISAFGNLRPLVLAEDAQRFLTALPFPKPKGNDNICODDLSITFTSPMSLDCLVVGSPRE- 800
 Db 756 PENG-----PVLDDSLPNSVHEIYIPAKDCGKKEKICISDLSHVATTEKDLIIIVRSQNDK 810

QY 801 FNVTVVRNDGDSYRTQVTPFPPLDLVSRKYSTLQNRQSRSRLACESASSTEVSGAL 860
 Db 811 FNVSLTVKTKDSAYNTRIVHYSNPLVSGIEAQKD-----SCSEN----- 853

QY 861 KSTSCSINHPIFPENSEVTFNITFDVDSKASIGN-KLLKANVTSENNMPTNKTFQLE 919
 Db 854 HNTCKVGYPFLRGEMVTKILFQNTSYLMENVTIYLSATSDSEEPETLSDNVVIS 913

QY 920 LPVKYAVMVVTSHGVSSTKYNFTASENTRVMOHQYQVSN-----LQORS-----L 966
 Db 914 IPKVEVGLQFTS-SASEYHISIAANETVPEVINSTEDIGNINIFYLIRKSGSPMPDEL 972

QY 967 PISLVF-----LVPVRLNQTVIMDRPQVTFSENLSSSTCHTKE-----RLPS 1007
 Db 973 KLSIGFPNMTNSGYPLVPTGLSS-----SENANCRPHIFEDPFSINSGRKMTT 1021

QY 1008 HSDFLAELRKAPVNCSTAVCORIQCDIPFFQIE-----EFNATLK 1049
 Db 1022 STD---HLKRGITLDCNTCKPATITCNLTSSDISQVNVSLILWKPTFIKSYFSSLNLTIR 1078

QY 1050 GNLSPDMWIKTSHNHLIVSTAEIILFNDISVFTLLPGQAFVRSQTBTKVPEVPEVNPPLP 1109
 Db 1079 GEL-----REENASLVLSNN-----QKRELAIQISKGLGRVPL 1114

QY 1110 ---IVGSSVGGHLLALITAAALYKLGFFKQYKDMSE 1144
 Db 1115 MWILLSAFAGLLLLILALWKIGFFRPLKKNMEK 1151

RESULT 8
 I45914
 integrin alpha 2 subunit - bovine (fragment)
 C:Species: Bos primigenius taurus (cattle)
 C>Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 15-Sep-2003
 C:Accession: I45914
 R:Kamata, T.; Puzon, W.; Takada, Y.
 J. Biol. Chem. 269, 9659-9663, 1994
 A:Title: Identification of putative ligand binding sites within the I-domain of integrin alpha 2 subunit.
 A:Reference number: A54402; MUID:94193647; PMID:7511592
 A:Accession: I45914
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-1170 <KAM>
 A:Cross-references: GB:I25886; NID:g439695; PIDN:AAB59255.1; PID:g439696
 F:161-336/Domain: von Willebrand factor type A repeat homology <WAA2>

Query Match 18.5%; Score 1099; DB 2; Length 1170;
 Best Local Similarity 27.8%; Pred. No. 9.6e-68;
 Matches 340; Conservative 215; Mismatches 498; Indels 168; Gaps 47;

QY 11 LTLCHGFNLDTENAMTFO-ENARGFQGSVVOL---QGSRVVVGAPQEIIVANORGLYQCD 66
 Db 13 LNCVAYNVGVLFPKAKIFGSPSSEQFGYAVQVQFINPKGNWLLVGSFWSGPFKNRGVDVYKC 72

QY 67 --DYSTGSCBPIRLQ-----VPEAVNMSLGLSLAATTPPQLLACGPTVHOTCSEN 117
 Db 73 PVDLSTTTCERKLI-QTSTSMNVMTKMTKSLGLTLFNVGCTGGFLTCGLPLWAQCGSQY 132

QY 118 YVKGCLFGLFGNLRQOPKPPALRGCPQEDSDIAFLDGGSGSIIPHFRMKPFVSTV 177
 Db 133 YTTGVCSDVSPDF-QLRTSFAPVQTCF-SFIDVWVWCDENSIYPMD--AVKNFLKFEV 188

F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-1163/Product: cell surface glycoprotein CD11c #status predicted <MAM>
F:20-1107/Domain: extracellular #status predicted <EXT>
F:149-319/Domain: von Willebrand factor type A repeat homology <VWA4>
F:1108-1133/Domain: transmembrane #status predicted <TM>
F:1134-1163/Domain: intracellular #status predicted <INT>
F:61,89,392,697,735,899,939,1050/Binding site: carbohydrate (Asn) (covalent) #status pre

Query Match 59.0%; Score 3513; DB 1; Length 1163;
Best Local Similarity 61.4%; Pred. No. 1,1e-235;
Matches 701; Conservative 138; Mismatches 296; Indels 6; Gaps 4;

QY 5 VLLTALTLCHGFNLDTENAMTFQENARGFQGVVYQLOQSRVVGAPQEIIVAAQNGSLY 64
DB 8 LLLTALATSLGFLNLTDELTAFVDSAGFGSDVVOYVANSVWVGAPQKITAANTGCLY 67
QY 65 QCDYSTGSCBRLQVPEAVNMSGLSLAATTSPLLACGPTVHOTCSENTYVKGCLP 124
DB 68 QCGYSTGACEPIGLQVPEAVNMSGLSLASTTSPLLACGPTVHCEGRNMYTLGLCF 127
QY 125 LFGSNLRQKQKPPALRGCPQEDSDIAPLIDGSGSIIPHDFRRMKEFVSTVMEQLKSK 184
DB 128 LLGPT-
QY 185 TLPSLMQYSEPRHFTKEFQNPNSRLVKPITQLLGRHTATGIRKRVRELNYNG 244
DB 186 TOFLMQSPNKPQHTFTPEERFRTSNPLSLASVHQLQGFTYTAIQQVVRHRLFHASYG 245
QY 245 ARKNAFKILVITDCEKFGDPLGVEDVPEADREGVIRYVIGVDAFSEKSRQELATIA 304
DB 246 ARRDATKILVITDCKEGSDLDYKDVIPMDADAAGIIRYALGVGLAFQNRNSWRELNDIA 305
QY 305 SKPRDRHVQNNFALXTIONLREKIPALHGTGTGSSSFHEHMSQEGSAITSNGP 364
DB 306 SKPSQEHFKVEDFDALXIOQNKKEKIPALHGTGTGSSSFHEHMSQEGSAITSNGP 365
QY 365 LLSTVGSVDWAGGVLYTSKEKSTFINTRVDSMDNDAYLGVAAAILRNVRQSLVLGAP 424
DB 366 VLGAQSVTSWGGAPLYPNKSPFINNSQENVDRDLSYLSYELALWGVQSVLVLGAP 425
QY 425 RYQHIGLVAMFRONTGMESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHY 484
DB 426 RYQHTGKAVITQVSRQWRMAEVTGTQIGSYFGASLCSVDVDTGSTDVLIGAPHY 485
QY 485 QTRGGVSVCPPLPRGORAWCDVLYGEOQPCWCFGAALTVLGDVNGDKLTDTVALGAP 544
DB 486 QTRGGVSVCPPLPRGWR-RWNCDAVLYGEOQPCWCFGAALTVLGDVNGDKLTDTVALGAP 544
QY 545 GEENRGAVYLFHGTSGSGISPSHSQRTAGSLSPRLQYFGQSLSGGQDLTQDGLAVDLAV 604
DB 545 GEENRGAVYLFHGTSGSGISPSHSQRTAGSLSPRLQYFGQSLSGGQDLTQDGLAVDLAV 604
QY 605 GAQGHVLLRSQPLRVKVAIMEENPREVARNVPCNDGVKKEAGEVRVCLHVQKSTRD 664
DB 605 GARGOVLRLTRPVLWGVSHQFTPAEIPRPAFCREQVSEQTLVGSNCLIKDKSKNS 664
QY 665 FLREGQIQSVVYTDIALDSGRPHRAVNETKSTRKQTVGLTQTCETLKLQLPNCIE 724
DB 665 LLGSRDLQSSVTLDLALDPGRSLSPRATFOETKRSLSRVRLGLKACENFNLPLPSQVE 724
QY 725 DPVSPVLRNFSVLGTPSLAFGNLRPVLAEDAQELTALPFPFKNGCNINICODLSIT 784
DB 725 DSVPTITLRLNFTLVGKPLAFRLNRLMLAADAQRYFTASLPFKNGCADHICODNLGIS 784
QY 785 FSEMSLCLVVGGRPREFNVTVTRNDGEDSVRTQVTFPPFLDLSYKRVSTLQNRORSW 844
DB 785 FSPFGLKSLVGSNLELNAEAVWVNDGEDSVYGTITFHPAGLSYRVVABQKQGLSL 844
QY 845 PLACESASTSVGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTS 904
DB 845 HLTCDSPVPG--SQGTWSTSCRNHLIPRGQAQITFLATFDVSPKAVLGDRLLTANVSS 902
QY 905 ENNMPTNKTEFOLELPVKYAVVWVTVSHGVSTKYLNFAS-ENTSRVMHQYQVSNLQ 963

DB 903 ENMPTNKTEFOLELPVKYAVVWVTVSHGVSTKYLNFSESEKESHVAMHRVQVNLQ 962
QY 964 RSLPISVLVLPVRLNQTVDWPRQVTFSENLSSTCHTKERLPSHSDFLAELKAPVNC 1023
DB 963 RDLFVSINFWVPELNOEAVMDVEVSHQPNQPSLRCSSEKIAAPPASDPLAHIQNPVLDC 1022
QY 1024 SIACVQRIQCDIPFGQIEBFNATLKGNSLSPDAVYIKTSHNHLIVSTAELFNDVSFTLL 1083
DB 1023 SIACVQRIQCDIPFGQIEBFNATLKGNSLSPDAVYIKTSHNHLIVSTAELFNDVSFTLL 1082
QY 1084 PGQAFVRSOTETKVPFVFNPLPLIFGVSSVGGLLLLALITAAKYKLGFPKQYKDMMS 1143
DB 1083 PGQAFVRSOTETKVPFVFNPLPLIFGVSSVGGLLLLALITAAKYKLGFPKQYKDMMS 1142
QY 1144 E 1144
DB 1143 E 1143

RESULT 4
S03308
N: cell surface glycoprotein CD11a precursor - human
C: Alternate names: leukocyte adhesion glycoprotein LFA-1 alpha chain; leukocyte function-
associated antigen 1
C: Species: Homo sapiens (man)
C: Date: 28-Feb-1990 #sequence, revision 28-Feb-1990 #text_change 20-Aug-1999
C: Accession: S03308; A47458; A47565; A48759; S36044
J: Larson, R.S.; Corbi, A.L.; Berman, L.; Springer, T.
J: Cell Biol. 108, 703-712, 1989
A: Title: Primary structure of the leukocyte function-associated molecule-1 alpha subunit
A: Reference number: A47458; MUID:93248261; PMID:8097887
A: Accession: S03308
A: Molecule type: mRNA
A: Residues: 1-1170 <LAR>
A: Cross-references: EMBL:Y00796; NID:g31421; PIDN:CAAG68747.1; PID:g31422
A: Note: part of this sequence was confirmed by protein sequencing
R: Cornwell, R.D.; Gollahan, K.A.; Hickstein, D.D.
Proc. Natl. Acad. Sci. U.S.A. 90, 4221-4225, 1993
A: Title: Description of the leukocyte function-associated antigen 1 (LFA-1 or CD11a) p.
A: Reference number: A47458; MUID:93248261; PMID:8097887
A: Accession: A47458
A: Molecule type: DNA
A: Residues: 1-20 <COR>
A: Note: sequence extracted from NCBI backbone (NCBI:130862, NCBI:130863)
R: Shelley, C.S.; Farokhzad, O.C.; Arnaout, M.A.
Proc. Natl. Acad. Sci. U.S.A. 90, 5384-5386, 1993
A: Title: Identification of cell-specific and developmentally regulated nuclear factors
A: Reference number: A47565; MUID:93281759; PMID:8099450
A: Accession: A47565
A: Molecule type: DNA
A: Residues: 1-20 <SHE>
A: Cross-references: GB:M95609
R: Nueda, A.; Lopez-Cabrera, M.; Vara, A.; Corbi, A.L.
J. Biol. Chem. 268, 19305-19311, 1993
A: Title: Characterization of the CD11a (alphaL, LFA-1alpha) integrin gene promoter.
A: Reference number: A48759; MUID:93374910; PMID:8103515
A: Accession: A48759
A: Molecule type: DNA
A: Residues: 1-20 <NUE>
A: Cross-references: EMBL:Z22804; NID:g311405; PIDN:CAA30461.1; PID:g311406
C: Genetics:
A: Gene: GDB:ITGAL; CD11A
A: Cross-references: GDB:119757; OMIM:153370
A: Map position: 16p11.2-16p11.2
C: Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hor
C: Keywords: cell adhesion; cytoskeleton; glycoprotein; heterodimer; surface antigen; ti
P: 1-25/Domain: signal sequence #status predicted <SIG>
F: 26-1170/Product: leukocyte adhesion glycoprotein LFA-1 alpha chain #status predicted
F: 154-317/Domain: von Willebrand factor type A repeat homology <VWA2>

Query Match 26.2%; Score 1559; DB 2; Length 1170;
Best Local Similarity 34.3%; Pred. No. 9.5e-100;
Matches 410; Conservative 210; Mismatches 464; Indels 110; Gaps 38;

A;Molecule type: DNA
A;Residues: 1-1153 <YTT>
A;Cross-references: EMBL:X07640; NID:g52982; PIDN:CAA30479.1; PID:g52983
A;Note: the authors translated the codon CAC for residue 569 as Gln
R;Sastre, L.; Roman, J.M.; Teplov, D.B.; Dreyer, W.J.; Gee, C.E.; Larson, R.S.; Roberts,
Proc. Natl. Acad. Sci. U.S.A. 83, 5644-5648, 1986
A;Title: A partial genomic DNA clone for the alpha subunit of the mouse complement recep
A;Reference number: 159078; MUID:86287312; PMID:2942940
A;Accession: 159078
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 11-44 <RES>
A;Cross-references: GB:M14293; NID:g198993; PIDN:AAA39484.1; PID:g554193
C;Genetics:
A;Gene: Mac-1
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
F;1-16/Domain: signal sequence #status predicted <SIG>
F;17-1153/Product: leukocyte surface glycoprotein Mac-1 alpha chain #status experimental
F;148-318/Domain: von Willebrand factor type A repeat homology <VWA2>
F;1106-1129/Domain: transmembrane #status predicted <TM>

Query Match 76.3%; Score 4545; DB 2; Length 1153;
Best Local Similarity 74.5%; Pred. No. 1.7e-307; Indels 2; Gaps 2;
Matches 860; Conservative 141; Mismatches 151;

Qy 1 MALRVLLLTALTLCHGFNLDTENAMTQENARGQSVQVQLQGSRRVVGAPQEIIVANQR 60
Db 1 MTLKALLVLTALCHGFNLDTENAMTQENARGQSVQVQLQGSRRVVGAPQEIIVANQR 60

Qy 61 GSLVQDYSGCSPIRLQVPEAVNNSGLSLAATSPOLLACGTHVQTSNTYVK 120
Db 61 GALYQCDYSTSRCHPIPLQVPEAVNNSGLSLAATSPOLLACGTHVQTSNTYVK 120

Qy 121 GLCLFGLSRLQOPKPEALRGCPQESDIAFLIDGSGSLIIPDFFRMKEFVSTVMEQL 180
Db 121 GLCLVFLSRLQOPKPEALRGCPQESDIAFLIDGSGSLIIPDFFRMKEFVSTVMEQL 180

Qy 181 KKSXTLSLMQYSBEPRIHFTFKFQNNPNRSLVXPKITQLLGHTHATGIRKVRRLFN 240
Db 181 KKSXTLSLMQYSBEPRIHFTFKFQNNPNRSLVXPKITQLLGHTHATGIRKVRRLFN 240

Qy 241 ITNCRANKAPKLVITIDGKFGPLGVEDVIPEADREGVIRYVIGVDAPRSKSKQEL 300
Db 241 ITNCRANKAPKLVITIDGKFGPLGVEDVIPEADREGVIRYVIGVDAPRSKSKQEL 300

Qy 301 NTIASKPRDRHVQVNNFEALKTIONQLEKIFAIEGTGTGSSSPHEMSQEGFSAAIT 360
Db 301 DTIASKPAGEHVQVNNFEALKTIONQLEKIFAIEGTGTGSSSPHEMSQEGFSAAIT 360

Qy 361 SNGILLSTVSGYDRAGGVPLTSKSTFTNMTVDSMDNDAYLGAAIILNRVQSLV 420
Db 361 SNGILLSTVSGYDRAGGVPLTSKSTFTNMTVDSMDNDAYLGAAIILNRVQSLV 420

Qy 421 LGAPRYCHIGLVAFRONTGKWSNANVKGTOIGAYFGASLCSDVDVDSNGSDTLVLICAP 480
Db 421 LGAPRYCHIGLVAFRONTGKWSNANVKGTOIGAYFGASLCSDVDVDSNGSDTLVLICAP 480

Qy 481 HYYTQTRGGQVSVCLPRGQBARQCDALVIGEGQPGWRGALTVLGDVNGDKLTIVA 540
Db 481 HYYTQTRGGQVSVCLPRGQBARQCDALVIGEGQPGWRGALTVLGDVNGDKLTIVA 540

Qy 541 IGAPGEDNRGAVLFEHTGSGSLSPSHSRIAGSKLSPRIQYFGQSLSGGQCDLTMDGLV 600
Db 541 IGAPGEDNRGAVLFEHTGSGSLSPSHSRIAGSKLSPRIQYFGQSLSGGQCDLTMDGLV 600

Qy 601 DLTVGAQCHVLLRSQVPLRVKALINEFNPREVARNFECNDQVVKGEAGEVRVCLHVQK 660
Db 601 DLTVGAQCHVLLRSQVPLRVKALINEFNPREVARNFECNDQVVKGEAGEVRVCLHVQK 660

Qy 661 STRDLREGIQSVVYDLDLSDGSPHSRAVNTKSTRTQVGLTQTCETKLQLP 720
Db 661 STRDLREGIQSVVYDLDLSDGSPHSRAVNTKSTRTQVGLTQTCETKLQLP 720

Qy 721 NCIEDPVSIVLRINPFLVGTPLSAGNLRPVLAEDAORLFTALPFPKNCNDNICQDD 780
Db 720 DCVDDSVSIIILRLNLTGRRPLRSGNLRPVLAEDAORLFTALPFPKNCNDNICQDD 779

Qy 781 LSTTFSFSLDCLVVGGRPEFNVTVVRNDGSDSYRTQVTFPPLDLSYRKVSTLQNR 840
Db 780 LSTTFSFSLDCLVVGGRPEFNVTVVRNDGSDSYRTQVTFPPLDLSYRKVSTLQNR 840

Qy 841 QRSWRL-ACESASSSTEVSGALSTSCSNHPIFPENSEVTENITPDVDSKASIGNKLK 899
Db 840 KKPWFVYKPAESSSSSGRGLKSTTNHPIFPANSEVTENITPDVDSHAGFGNKLK 899

Qy 900 ANVTSENNMPTNKTEFQLELFPKAVVMVTVSHGVSTKYLNFNTASENTRVYHQYQVS 959
Db 900 ALVASENNMPTNKTEFQLELFPKAVVMVTVSHGVSTKYLNFNTASENTRVYHQYQVS 959

Qy 960 NLGQSLPVLVPLVRLNQVTVDRPQVTSSENLSSCTKTKERLPSHSDFLAELKAP 1019
Db 960 NLGQSLPVLVPLVRLNQVTVDRPQVTSSENLSSCTKTKERLPSHSDFLAELKAP 1019

Qy 1020 VVNCISLAVCORIQCDIPFPGIOEFNATLKGNSPDWYIKTSHNHLIIVSTABILEFNSV 1079
Db 1020 VVNCISLAVCORIQCDIPFPGIOEFNATLKGNSPDWYIKTSHNHLIIVSTABILEFNSV 1079

Qy 1080 FTLLPQCGAFVRSQTSKTEKPEFVNPPLIVGSSVGGLLALITAAALYKLGFFKQYK 1139
Db 1080 FTLLPQCGAFVRSQTSKTEKPEFVNPPLIVGSSVGGLLALITAAALYKLGFFKQYK 1139

Qy 1140 DMKSEGGPPCARPQ 1153
Db 1140 DMKSEGGPPCARPQ 1153

RESULT 3

RWHLIC

Cell surface glycoprotein CD11c precursor - human

N;Alternate names: leukocyte adhesion receptor p150,95 alpha chain

C;Species: Homo sapiens (man)

C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 22-Jun-1999

C;Accession: A36584; A35543; S00864

R;Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.

J. Biol. Chem. 265, 12750-12751, 1990

A;Reference number: A36584

A;Contents: erratum

A;Accession: A36584

A;Molecule type: DNA

A;Residues: 1-1163 <COR>

A;Note: this revision to the sequence from reference A35543 includes the carboxyl end

R;Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.

J. Biol. Chem. 265, 2782-2788, 1990

A;Title: Genomic structure of an integrin alpha subunit, the leukocyte p150,95 molecule

A;Reference number: A35543; MUID:90153906; PMID:2303425

A;Accession: A35543

A;Molecule type: DNA

A;Residues: 1-834 <CO2>

A;Note: this sequence has been revised in reference A36584

R;Corbi, A.L.; Miller, L.J.; O'Connor, K.; Larson, R.S.; Springer, T.A.

EMBO J. 6, 4023-4026, 1987

A;Title: cDNA cloning and complete primary structure of the alpha subunit of a leukocyte

A;Reference number: S00864; MUID:88166645; PMID:3327687

A;Accession: S00864

A;Molecule type: mRNA

A;Residues: 1-755, 'L', 757-1163 <CO3>

A;Cross-references: GB:M81695; EMBL:Y00093; NID:g487829; PIDN:AAA59180.1; PID:g487830

A;Note: part of this sequence was confirmed by protein sequencing

C;Comment: A common beta chain (CD18) forms a heterodimer with CD11b to form Mac-1 on n

C;Genetics:

A;Gene: ITGA1; CD11C

A;Cross-references: GDB:119758; OMIM:151510

A;Map position: 16p11.2-16p11.2

C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom

C;Keywords: calcium; cell adhesion; glycoprotein; heterodimer; magnesium; tandem repeat

A>Note: part of this sequence was confirmed by protein sequencing
 R:Fleming, J.C.; Pahl, H.L.; Gonzalez, D.A.; Smith, T.F.; Tenen, D.G.
 J. Immunol. 150, 480-490, 1993
 A:Title: Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-1
 n during evolution.
 A:Reference number: A46526; MUID:93123748; PMID:8419480
 A:Accession: A46526
 A:Status: not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-499,501-1153 <FILE>
 A:Cross-references: GB:S52227; PIDN:AAB24821.1; PID:9263049
 A>Note: the last three bases of intron 13, CAG, are included in some but not all mature
 A:Note: sequence extracted from NCBI backbone (NCBI:P121963)
 R:Pierce, M.W.; Remold-O'Donnell, E.; Todd III, R.F.; Arnaout, M.A.
 Biochim. Biophys. Acta 874, 368-371, 1986
 A:Title: N-terminal sequence of human leukocyte glycoprotein Mol: conservation across sp
 A:Reference number: A90664; MUID:87076671; PMID:3539202
 A:Accession: A26091
 A:Molecule type: protein
 A:Residues: 17-31 <PIE>
 A:Experimental source: granulocytes
 R:Pahl, H.L.; Rosmarin, A.G.; Tenen, D.G.
 Blood 79, 865-870, 1992
 A:Title: Characterization of the myeloid-specific CD11b promoter.
 A:Reference number: I52567; MUID:92144986; PMID:1346576
 A:Accession: I52567
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-9 <RES>
 A:Cross-references: GB:M84477; NID:9180184; PIDN:AAA51960.1; PID:9553219
 A:Comment: A common beta chain (CD18) forms a heterodimer with this chain to form Mac-1
 C:Genetics:
 A:Gene: GDB:ITGAM; CR3A
 A:Cross-references: GDB:120599; OMIM:120980
 A:Map position: 16p11.2-16p11.2
 A:Note: promoter contains a GATA motif and two Sp1 consensus binding sites
 C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
 C:Keywords: alternative splicing; calcium; cell adhesion; glycoprotein; heterodimer; mag
 F:1-16/Domain: signal sequence #status predicted <SIG>
 F:17-1153/Product: cell surface glycoprotein CD11b #status experimental <MAT>
 F:17-1108/Domain: extracellular #status predicted <EXT>
 F:148-318/Domain: von Willebrand factor type A repeat homology <WA2>
 F:465-478/Region: calcium/magnesium binding #status predicted
 F:530-538/Region: calcium/magnesium binding #status predicted
 F:593-601/Region: calcium/magnesium binding #status predicted
 F:1109-1134/Domain: transmembrane #status predicted <TM>
 F:1135-1153/Domain: intracellular #status predicted <INT>
 P:86.240.391,469,693,697,735,802,881,901,912,941,947,979,994,1022,1045,1051,1076/Binding

Query Match 100.0%; Score 5956; DB 1; Length 1153;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRVLLLTALTLCGPNLDTENAMTFQENARGFGQSVVQLQSSRVVVGAPQEIIVANQR 60
 DB 1 MALRVLLLTALTLCGPNLDTENAMTFQENARGFGQSVVQLQSSRVVVGAPQEIIVANQR 60

QY 61 GSLYQCDYSTGSCPEIRIQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVK 120
 DB 61 GSLYQCDYSTGSCPEIRIQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVK 120

QY 121 GLCLFPGSNLRQOPKPFQALRGCPQEDSDIAFLDGSGLIIPHDPRMKFEVSTVMEQL 180
 DB 121 GLCLFPGSNLRQOPKPFQALRGCPQEDSDIAFLDGSGLIIPHDPRMKFEVSTVMEQL 180

QY 181 KKSXTLFLSMQYSEEFRIHTFKFQNNPNRSLVKPTITQLLGRTHATGIRKVVRLFN 240
 DB 181 KKSXTLFLSMQYSEEFRIHTFKFQNNPNRSLVKPTITQLLGRTHATGIRKVVRLFN 240

QY 241 ITNGARKNAFKILVITDGEKFGPLGYEDVIPADREGVIRYVIGVDAPFRSEKSRQEL 300
 DB 241 ITNGARKNAFKILVITDGEKFGPLGYEDVIPADREGVIRYVIGVDAPFRSEKSRQEL 300

QY 301 NTIASKPRDRHVQVNNFEALKTIONLRKIFAIETGTGSSSSSFEHMSQEGFSAIT 360
 DB 301 NTIASKPRDRHVQVNNFEALKTIONLRKIFAIETGTGSSSSSFEHMSQEGFSAIT 360

QY 361 SNGPILLSTVGSYDWAGGVFLYTSKEKSTFINMTVRVDSMDNDAYLGAAAIIILNRVQSLV 420
 DB 361 SNGPILLSTVGSYDWAGGVFLYTSKEKSTFINMTVRVDSMDNDAYLGAAAIIILNRVQSLV 420

QY 421 LGAPRYOHIGLVAMFRONTGMWESNANVKGTQICAYFGASLCSDVDNSGSTDILVLIGAP 480
 DB 421 LGAPRYOHIGLVAMFRONTGMWESNANVKGTQICAYFGASLCSDVDNSGSTDILVLIGAP 480

QY 481 HYYOTRGQGVSVCPPLPRGQARWQCDVLYGEGQPGWRFGAALTIVLGDVNGDKLTDVA 540
 DB 481 HYYOTRGQGVSVCPPLPRGQARWQCDVLYGEGQPGWRFGAALTIVLGDVNGDKLTDVA 540

QY 541 IGAPGEDNNGAVLPHGTSGSLSPSHSORIAGSKLSPLOYEGOSLGGODLTMGLV 600
 DB 541 IGAPGEDNNGAVLPHGTSGSLSPSHSORIAGSKLSPLOYEGOSLGGODLTMGLV 600

QY 601 DLTVGAQGHVLLLSQPLRVKAIMFNPREVARNVFECDNVVYKGEAGEVRVCLHVQK 660
 DB 601 DLTVGAQGHVLLLSQPLRVKAIMFNPREVARNVFECDNVVYKGEAGEVRVCLHVQK 660

QY 661 STRDLREGQIOSVVTYDLALDSGRPHSRVAFNETKSTRTROTQVGLTTCETLKLQLP 720
 DB 661 STRDLREGQIOSVVTYDLALDSGRPHSRVAFNETKSTRTROTQVGLTTCETLKLQLP 720

QY 721 NCIEDPVPVILRLNPSLVGTPLSAFGNLRPLVAEDAQRPLTALFPPEKNCNDNICODD 780
 DB 721 NCIEDPVPVILRLNPSLVGTPLSAFGNLRPLVAEDAQRPLTALFPPEKNCNDNICODD 780

QY 781 LSITFSFMSLDCLVWGPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNS 840
 DB 781 LSITFSFMSLDCLVWGPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNS 840

QY 841 QRSWRLACESASSSTEVSGLKSTSCSINHIFIPFENSEVTENITFDVDSKASLGKLLKA 900
 DB 841 QRSWRLACESASSSTEVSGLKSTSCSINHIFIPFENSEVTENITFDVDSKASLGKLLKA 900

QY 901 NVTSENMPRTNTEPOLELPVKYAVVMVTSHGVSSTKYLNFTASENTSRVMHQYQVSN 960
 DB 901 NVTSENMPRTNTEPOLELPVKYAVVMVTSHGVSSTKYLNFTASENTSRVMHQYQVSN 960

QY 961 LGQSLPISLVLPVVRINQTVWDRPQVTFNSNLSTCHTKERLPSHSDFLAELKAPV 1020
 DB 961 LGQSLPISLVLPVVRINQTVWDRPQVTFNSNLSTCHTKERLPSHSDFLAELKAPV 1020

QY 1021 VNCIAVCQRIQCDIPFGIOBEFNATLKGNSLFDWYIKTSHNHLIIVSTABILFNDVSF 1080
 DB 1021 VNCIAVCQRIQCDIPFGIOBEFNATLKGNSLFDWYIKTSHNHLIIVSTABILFNDVSF 1080

QY 1081 TLLPGQAGAPVRSOTETKVEPFEVNPPLPLTVGSSVGLLALLIITAALYKLGFKQYKD 1140
 DB 1081 TLLPGQAGAPVRSOTETKVEPFEVNPPLPLTVGSSVGLLALLIITAALYKLGFKQYKD 1140

QY 1141 MMSEGGPPGAEPPQ 1153
 DB 1141 MMSEGGPPGAEPPQ 1153

RESULT 2

S00551

leukocyte surface glycoprotein Mac-1 alpha chain precursor - mouse

N;Alternate names: complement-3 receptor alpha chain

C;Species: Mus musculus (house mouse)

C;Date: 30-Sep-1989 #sequence_revision 30-Sep-1991 #text_change 22-Oct-1999

C;Accession: S00551; I59078

R;Pyteia, R.

EMBO J. 7, 1371-1378, 1988

A;Title: Amino acid sequence of the murine Mac-1 alpha chain reveals homology with the

A;Reference number: S00551; MUID:88312584; PMID:3044779

A;Accession: S00551

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 17:05:16 ; Search time 14.7639 Seconds
(without alignments)
7512.163 Million cell updates/sec

Title: US-09-902-481b-1

Perfect score: 5956

Sequence: 1 MALRVLLTALTLCHGFNLD.....FTKQYKDMSEGGPPGAEFQ 1153

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: pir1.*
- 2: pir2.*
- 3: pir3.*
- 4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	5956	100.0	1153	1 RWHULB	cell surface glyco
2	4545	76.3	1153	2 S00551	leukocyte surface
3	3513	59.0	1153	1 RWHULC	cell surface glyco
4	1559	26.2	1170	2 S03308	cell surface glyco
5	1549	26.0	1163	2 I56126	lymphocyte function
6	1163	19.5	1179	2 A53213	integrin alpha-E c
7	1108.5	18.6	1151	2 A45226	integrin alpha-1 c
8	1099	18.5	1170	2 I45314	integrin alpha 2 s
9	1089	18.3	1178	2 S44142	VIa-2 protein homo
10	1087.5	18.3	1181	2 A33998	integrin alpha-2 c
11	1082	18.2	1180	2 A35854	integrin alpha-1 c
12	667	11.2	1039	2 A41131	lymphocyte-Peyer's
13	638	10.7	1038	2 S06046	integrin alpha-4 c
14	630	10.6	1035	2 I58409	integrin alpha-9 c
15	614.5	10.3	1041	2 T31437	integrin alpha cha
16	579.5	9.7	1054	2 JC7294	alnap integrin -
17	572.5	9.6	1051	2 A35761	cell surface glyco
18	567.5	9.5	1053	2 I55534	VIa-3 alpha subuni
19	555.5	9.3	1053	2 S44250	integrin alpha-5 c
20	550	9.2	1034	2 A36108	integrin alpha-v c
21	535	9.0	1044	2 T10050	integrin alpha-v c
22	532	8.9	1049	2 A27079	fibronectin recept
23	532	8.9	1073	2 B36429	integrin alpha-6 c
24	530.5	8.9	1072	2 A48457	integrin alpha-6 c
25	529.5	8.9	1051	2 A40021	integrin alpha-3 c
26	526	8.8	1048	2 A27421	integrin alpha-5 c
27	525.5	8.8	1091	2 A41543	integrin alpha-6 c
28	517	8.7	1044	2 S16516	integrin alpha-8 c
29	505	8.5	1394	2 A29637	position-specific

30	496.5	8.3	1146	2 S40311	integrin - fruit f
31	495.5	8.3	1039	2 A34269	integrin alpha-2b
32	492.5	8.3	1037	2 A60163	glycoprotein Iib -
33	490	8.2	1137	2 JC5950	integrin alpha-7 c
34	488	8.2	1135	2 I61186	alpha-7 integrin -
35	486	8.2	126	2 B30892	leukocyte adhesion
36	474.5	8.0	1226	2 S44824	P5472.1 protein -
37	469.5	7.9	1106	2 S38783	integrin alpha cha
38	454.5	7.6	1139	2 S28277	hypothetical prote
39	454	7.6	1045	2 S60571	integrin alpha v c
40	430.5	7.2	1115	2 T09433	integrin alpha cha
41	423.5	7.1	1115	2 T09403	integrin alpha cha
42	391	6.6	764	2 I36916	glycoprotein Iib -
43	309	5.2	1086	2 T18523	integrin alpha cha
44	304.5	5.1	272	2 A55348	integrin alpha-1 -
45	299	5.0	604	2 I36917	glycoprotein Iib -

ALIGNMENTS

RESULT 1

RWHULB

cell surface glycoprotein CD11b precursor [validated] - human
N;Alternate names: complement receptor type 3 alpha chain; leukocyte adhesion protein 1
eukocyte integrin alpha chain; neutrophil adherence receptor alphaM chain
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text change 31-Dec-2000
C;Accession: A31108; A28915; A41600; A30892; A32218; A46526; A26091; I52567
R;Corbi, A.L.; Kishimoto, T.K.; Miller, L.J.; Springer, T.A.
J Biol. Chem. 263, 12403-12411, 1988
A;Title: The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, C
B.
A;Reference number: A31108; MUID:88315033; PMID:2457584
A;Accession: A31108
A;Molecule type: mRNA
A;Residues: 1-1153 <COR>
A;Cross-references: GB:J03925; NID:gi87284; PIDN:AAA59544.1; PID:G307148
A;Note: part of this sequence was confirmed by protein sequencing
R;Arnaut, M.A.; Gupta, S.K.; Pierce, M.W.; Tenen, D.G.
J. Cell Biol. 106, 2153-2158, 1988
A;Title: Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor
A;Reference number: A28915; MUID:88257215; PMID:2454931
A;Accession: A28915
A;Molecule type: mRNA
A;Residues: 1-499,501-965, 'P', 967-1153 <ARN>
A;Cross-references: GB:M18044; GB:J03270; GB:M19664; GB:X07421; NID:gi86935; PIDN:AAA5;
A;Note: the authors translated the codon TAC for residue 1129 as Thr
A;Note: part of this sequence, including the amino end of the mature protein, was confi
R;Shelley, C.S.; Arnaut, M.A.
Proc. Natl. Acad. Sci. U.S.A. 88, 10525-10529, 1991
A;Title: The promoter of the CD11b gene directs myeloid-specific and developmentally re
A;Reference number: A41600; MUID:92073318; PMID:1683702
A;Accession: A41600
A;Molecule type: DNA
A;Residues: 1-9 <SHE>
A;Cross-references: GB:M76724; NID:gi180018; PIDN:AAA58410.1; PID:G553215
R;Arnaut, M.A.; Remold-O'Donnell, E.; Pierce, M.W.; Harris, P.; Tenen, D.G.
Proc. Natl. Acad. Sci. U.S.A. 85, 2776-2780, 1988
A;Title: Molecular cloning of the alpha-subunit of human and guinea pig leukocyte adhe
A;Reference number: A34193; MUID:86190151; PMID:2833753
A;Accession: A30892
A;Molecule type: mRNA
A;Residues: 917-1042 <AR2>
A;Cross-references: GB:M18044
R;Hickstein, D.D.; Hickey, M.J.; Ozols, J.; Baker, D.M.; Back, A.L.; Roth, G.J.
Proc. Natl. Acad. Sci. U.S.A. 86, 257-261, 1989
A;Title: cDNA sequence for the alphaM subunit of the human neutrophil adherence recept
A;Reference number: A32218; MUID:8909893; PMID:2563162
A;Accession: A32218
A;Molecule type: mRNA
A;Residues: 9-1153 <HIC>
A;Cross-references: GB:J04145; NID:gi89068; PIDN:AAA59903.1; PID:G386975

Qy 721 SLVGTPLSAGNLPVLAEDAOPLTALPPEKNCNDNICODDLSITPSPASLDCLVG 780
 Db 737 TLVGKPLAFNLRLPMLAALQRIFTASLPPEKNCADHICODMIGISFSPGUKSLVG 796
 Qy 781 GPREFNVTVRNDGDSYRTQVTFPPFLDLSYKVKSTLQNRQSRWRLACESASTEV 840
 Db 797 SNLEBAEVMWMDGDSYGTITFESHAPGLSYRYVAEGQKQGLRSLHLTC--CSAPVG 854
 Qy 841 SGALKSTSCSINHDIPEENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEP 900
 Db 855 SQGTWSTSCRINHLIFRGAGITELATFDVSPKAVGLDRLLLIANVSSENNIPRTSTIF 914
 Qy 901 QLELPVYAVVYVTSYSHGVSTKYLNFTAS-ENTSRVMQHOYQVSNLQORSILPTSLVLP 959
 Db 915 QLELPVYAVVYVSSHEQFTKYNLFSESEKESHVAMHRVQVNNLQORDLPVSNFWP 974
 Qy 960 VRLNQTWIDPEQVTPSENLSSTCHYERLPSPHSDFTLAELKAPVNCSTAVOQIOCDI 1019
 Db 975 VELNOEAVMDVESHFQNPSPSRCSSEKIAPPASDFLAHQKPNVLPDCSIAGCLRFCDV 1034
 Qy 1020 PFGIOEENATLKNLSFDWYIKTSHNHLIVSTAILFNDSVFTLLPGQAFVRSOTE 1079
 Db 1035 PSFSQBELDFTLKNLSFGWROILOKYVSUSVAILIIPDTSYSQLPQEAFAWRAQTI 1094
 Qy 1080 TKVEPFPVPLPLIVGSSVGGLLLLALITAAALYKLGFFKRYKQYKDMNSE 1128
 Db 1095 TVLEKYVHNPILIVGSSIGGLLLALITAVLYKVGFFKRYKQYKEMNEE 1143

RESULT 14
 US-09-350-259-2
 ; Sequence 2, Application US/09350259
 ; Patent No. US20020062008A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Gallatin, Michael W.
 ; APPLICANT: Van der Vlieten, Monica
 ; TITLE OF INVENTION: No. US20020062008A1 Human 2
 ; FILE REFERENCE: 27866/35004
 ; CURRENT APPLICATION NUMBER: US/09/350, 259
 ; CURRENT FILING DATE: 1999-07-08
 ; EARLIER APPLICATION NUMBER: 09/193,043
 ; EARLIER FILING DATE: 1998-11-16
 ; EARLIER APPLICATION NUMBER: 08/173,497
 ; EARLIER FILING DATE: 1993-12-23
 ; EARLIER APPLICATION NUMBER: 08/286,889
 ; EARLIER FILING DATE: 1994-08-05
 ; EARLIER APPLICATION NUMBER: 08/362,652
 ; EARLIER FILING DATE: 1994-12-21
 ; EARLIER APPLICATION NUMBER: 08/943,363
 ; EARLIER FILING DATE: 1997-10-03
 ; NUMBER OF SEQ ID NOS: 114
 ; SOFTWARE: Patent In Ver. 2.0
 ; SEQ ID NO 2
 ; LENGTH: 1161
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-350-259-2

Query Match 57.6%; Score 3388; DB 9; Length 1161;
 Best Local Similarity 58.9%; Pred. No. 5e-304;
 Matches 665; Conservative 168; Mismatches 288; Indels 8; Gaps 6;

Qy 1 FNLDTENAMTFQENARFGQSVQLOGSRVVVGAPOIIVAAORGSLYQCDYSTGSCPEI 60
 Db 17 FNLDVEBPTIFQEDAGFGQSVQVQFSGSLVVGAPLEVVAANTGRLYDCAATGMCQPI 76
 Qy 61 RLQVFEAVNMSGLSLAATSPOLLACQPTHQTCSENTYKGLCFLFGSNLRQOPQK 120
 Db 77 PLHIRPEAVNMSGLTLAATNGSRLACQPTLHRVCGENSYSKGSCLLIGSRW-ETIQT 135
 Qy 121 FPRALRCPOEDSDIAPLDGSSIIIPHDFRMEKLVSTIMEOLKSKTFLSLMOYSEEP 180

Db 136 VPDATECPEQEMDIIVPLIDGSGSIDQNDNFQNMKGFPVQAVMGOFECTDTLIPALMOYSNLL 195
 Qy 181 RIHFTYKBFONNPNPSLIKPITOLLGRTHATGLKRYVRELFINITNGARKNAFKILFLL 240
 Db 196 KIHFTTQRTSPSOQSLVDPIVOLKGLTFTATGILTVVTLPHKNGARKSAKILIVI 255
 Qy 241 TDCEKGDPLGYBDVIPELDREGVIRYVLGFGDAPRSEKSRQELNLTVAKPPPDHVFQAN 300
 Db 256 TDQCKYDPLEYSDVIPAQEKAGIRYAVGVGHAFQPTARQELNTISAPPDHVFQVD 315
 Qy 301 NFALTKVQVQAREKIPATEGTQSGSSPHEHMSQEGSAAITSNGLPILSTVSGVDWAG 360
 Db 316 NFAALGSIQKLOEKIYAVEGTQSRASSPHEHMSQEGFSTALTMDELFLGAVGSSWSG 375
 Qy 361 GVFLXTSKSTPINTRVDSMDNDAYLGVAAAILRNVRQSLVAGAPRYQHIQLVAMER 420
 Db 376 GAFYPPNWSPPFINXQENVDWROSYLGYSTELALWKGVNLVGLAPRYQHTKAVIFT 435
 Qy 421 QNTGWMESNANVKGTOIGAYFGASLCSVDVDSNGSDTLVLIGAPHYYEOTRGQSVCPPL 480
 Db 436 QVSRQWRKKAETGTQIGSYFGASLCSVDVDSNGSDTLVLIGAPHYYEOTRGQSVCPPL 495
 Qy 481 PRGORAWOCDVLYGCEQCPKGRFGAALTVLGDVNGDKLTQVAIGARGEEDNRGAVYLF 540
 Db 496 PRQVRVQWQCDVLRGEQHPGRFGAALTVLGDVNGDKLTQVAIGARGEEDNRGAVYLF 555
 Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMGDLVDTGAQGHVLLRSQ 600
 Db 556 HGASBSGSIPSHSQRIASSOLSPLQYFGQALSGGQDLTMGDLVAVGARGQVLLRSL 615
 Qy 601 PVLKVAIMEFNPREVARNVPCNDQVVKKEAGVYVCLHWQKSTEDLAREQIOISVYT 660
 Db 616 PVLKVGVMRFPSEVEKAVYRCWEKPSALEAGDATVCLITQKSLDQI--GDIQSSVR 673
 Qy 661 YDLALDGRPHSRVAVENETKNSRRTQVLTGTCTETLKLQLPNCIEBPVSVIVRLNF 720
 Db 674 FDLALDGRPLTSRAIFNETKNPLTRKTLGLGHCHCETLKLPCDVEDVSPILHLNF 733
 Qy 721 SLVGTPLSAGNLPVLAEDAOPLTALPPEKNCNDNICODDLSITPSPASLDCLVG 780
 Db 734 SLVREPIPSQNLRLPVLAVGSDLPFTASLPPEKNCQDGLCEGDLGVTLSFSGLTITVG 793
 Qy 781 GPREFNVTVRNDGDSYRTQVTFPPFLDLSYKVKSTLQNRQSRWRLACESASTEV 840
 Db 794 SSELNVIUVNAGDSYGTVVSYPAGLSHRRVSGAQKQPHQSALLACETV-PTED 852
 Qy 841 SGALKSTSCSINHDIPEENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEP 900
 Db 853 EG-LRSSRCVNHPIPHGSGNGTFIVTFDVSYKATIGDRMLARASSENKASSKATP 911
 Qy 901 QLELPVYAVVYVTSYSHGVSTKYLNFTAS-ENTSRVMQHOYQVSNLQORSILPTSLVLP 959
 Db 912 QLELPVYAVVYVTSYSHGVSTKYLNFTAS-ENTSRVMQHOYQVSNLQORSILPTSLVLP 971
 Qy 960 VRLNQTWIDPEQVTPSENLSSTCHYERLPSPHSDFTLAELKAPVNCSTAVOQIOCDI 1019
 Db 972 VLLNGVAVMDVMEAPSSQSL--PCVSEKPPQHSDFLTQISRSPLMDCSIADCLQPRCDV 1029
 Qy 1020 PFGIOEENATLKNLSFDWYIKTSHNHLIVSTAILFNDSVFTLLPGQAFVRSOTE 1079
 Db 1030 PSFSQBELDFTLKNLSFGWRETLQKVLVVSVAEITPDTSVYSQLPQEAFAWRAQME 1089
 Qy 1080 TKVEPFEVENPLPLIVGSSVGGLLLLALITAAALYKLGFFKRYKQYKDMNSE 1128
 Db 1090 MYLEDEEVNAPILINGSSVGGALLLALITATLYKLGFFKRYKEMLEED 1138

RESULT 15
 US-09-891-943-2
 ; Sequence 2, Application US/09891943
 ; Publication No. US20030077278A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Gallatin, Michael W.

FT	CARBOHYD	980	980	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	1045	1045	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	1055	1055	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	1074	1074	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	1085	1085	N-LINKED (GLCNAC. . .) (POTENTIAL).	
SQ	SEQUENCE	1151	AA; 127837 MW; 693F3C1AABF52808 CRC64;		
Query Match					18.6%; Score 1108.5; DB 1; Length 1151;
Best Local Similarity					28.0%; Pred. No. 1.1e-65;
Matches					346; Conservative 210; Mismatches 486; Indels 195; Gaps 44;
QY	17	FNLDTENAMTPOENARG-FCQSVQL---	CGSRVVVGAPOEIVAAQNRGLYQCDISTGS	72	
DB	1	FNVGVKSMTPSGVEDMFGYVQYENEBGKWLGLSGPLVGQPKRTGDVYKCPVGRGE	60		
QY	73	CEP-IRLQVPEA-----	VMSLGLSLAATTSPOLLACGPTVHQTCSENVYKGL	122	
DB	61	SLPCVKLDLFPNTSIPNVTEVENMTFGSTL-VTPNPGFLACGLYAYECGHLHTYGI	119		
QY	123	CFLFGSLRQOPKFPALRGCPQEDSDIAPLDGSGSIIPHPFRMKFVSTVMBQLK-	181		
DB	120	CSVSPTFQVNSIAP--VOECSTQ-LDVIIVLDGNSIYPWD--SVTAPLNDLLKRMDI	174		
QY	182	-KSKTLFSLMOYSEEFRIHTEFEFONNPNRSLVXPITOLLGR-THTATGIRKVVRELF	239		
DB	175	GPQQTQVGIQVYGENVTHEFNANKYSTEEVIVAAKIIVQGRQWTALGTDTARKEAF	234		
QY	240	NIINGARONAKILVITDGEKFDPLGYEDVTPADREGVIRYIVGVGDAFR-----GE	294		
DB	235	TEARGARGVKVWVITDGESH-DNHRLLKKVQDCEDENIQRFSAIILGSYNRGNLSYE	293		
QY	295	XSQELNTIASKPPRPHVFOVNFELAKTIQNLQREKIPIAIEGTQGGSSSFHEMSQEG	354		
DB	294	KFVEEIKSIASETEKHEFFNVSDLLALVTIVKLGRIFALEATADQSAASFEMNSQTG	353		
QY	355	FSAAITSNGELLSTVSGSYDWAGGVFLYTSXB-----KSTF-INMTRVDSMDNDAYLGAA	408		
DB	354	FSAHYSQDWMLGAVGAYDNWGTWQKASQIIIPRNETFNVESTKKNRPL-ASYLGYTV	412		
QY	409	AIIILNRVQSL-VLGPAPRIYOHIGLVAMFRONTGWESNANVKTOIGAYFGASLCSVDVD	467		
DB	413	NSATASGVDLYTAGOPRYNHTGVIYRMEDGNIKILOLQSGEYSGVFSGLITTTDID	472		
QY	468	SNGSTDLVLTGAPHY-----	YEQTR--GGQVSVCLPRGQARWQCDVAVLY	511	
DB	473	KDSNTDILLGAPWNGTEKEBQGVVYVALNQTRFEYQMSLEPIKQTCSSRQHSNCTT	532		
QY	512	GEGQQPWG-RFGAALTVLGVNGKLTDAIGAPEGEDNRGAYLVFHTGSGGISPSHSQ	570		
DB	533	ENKNEPCGARFGTAIAAVKDLNLDFENDIVIGAPLEDHGGAYIYTHG-SGKTIKREYAQ	591		
QY	571	RIAGSKLSPLQVFGQSLGGQBLTMDGLVLDLVGAQGHVLLRLSOPVLRVKALMEFNR	630		
DB	592	RIPSGGDKTLKFPQASIHGEMDLNGDLTDVTIIGLGAALFWSRDVAVKVTWTFEN	651		
QY	631	EVARNPECDNQVVKGEAG--EVRVCLHVQ-KSTEDRLREGQIQSVVTDLALDSGRPH	687		
DB	652	KVNIQKKNCH--MEGKETVCINATVCFEVLKSKEDTIVEADLQ-----YRVTLDSLROI	704		
QY	688	SRAVENET-----KNSRQTOVGLTQTCETLKLQNLNCIEDPSPVIVLRALNESLVGT	741		
DB	705	SRSPFGSGTQERKVORNIIVRKSSC-----TKHSFYMLDXHDFQDSVYR---ITLDFNLT-D	755		
QY	742	PLSAFGLNRLPVLAEADQRLTALFPPEFKXGNDNICQDLSITFFSFWSLDCLVVGQPRE-	800		
DB	756	PENG-----PVLDSLPNSVHEIYIPAKOCGNKEKICISLSLHVATTEKDLIVRSQNDK	810		
QY	801	FNVTVVRNDGDSYETQVTFPPLDLSYRKVSTLQNRQSRWRLACESASTEVSGAL	860		
DB	811	FNVSLTVKTKDQAYNTRITVYHSPNLVPSGIEAIQKD-----SCSEN-----	853		
QY	861	KSTSCSINRPIPFENSEVTFNITFDVDSKASLGN-KLLKANVTSENMPRNTKTEFOLE	919		

RESULT 10

ITAH_HUMAN

ID ITAH_HUMAN STANDARD; PRT; 1189 AA.

AC Q9UKX5; Q9UKQ1;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DB Integrin alpha-11 precursor.

GN ITG11

OS Homo sapiens (Human)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI TaxID:9606;

RN [1]

SEQUENCE FROM N.A.

RC TISSUE=Petal heart, and Osteoblast;

RC MEDLINE=99417678; PubMed=10486209;

RA Lehner K., Ni J., Leung E., Gough S.M., Weaver A., Yao W.P., Liu D.,

Wang S.-X., Morris C.M., Krissansen G.W.;

"Cloning, sequence analysis, and chromosomal localization of the novel

human integrin alpha11 subunit (ITG11).";

Genomics 60:179-187(1999).

RL [2]

SEQUENCE FROM N.A.

RC TISSUE=Fetal muscle, and Uterus;

RC MEDLINE=99395147; PubMed=10464311;

RA Velling T., Kusche-Gullberg M., Sejersen T., Gullberg D.;

"cDNA Cloning and Chromosomal Localization of Human alpha(11)

Integrin. A collagen-binding, i domain-containing, beta(1)-associated

integrin alpha-chain present in muscle tissues.";

J. Biol. Chem. 274:25735-25742(1999).

RN [3]

SEQUENCE OF 954-1188 FROM N.A.

RC TISSUE=Fibroblast;

RA Andreu N., Estivill X., Escarceller M., Sumoy L.;

Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.

CC FUNCTION: INTEGRIN ALPHA-11/BETA-1 IS A RECEPTOR FOR COLLAGEN.

CC SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-11

ASSOCIATES WITH BETA-1.

CC SUBCELLULAR LOCATION: Type I membrane protein.

CC TISSUE SPECIFICITY: ACCORDING REF.2 HIGHEST LEVELS IN UTERUS AND

HEART, INTERMEDIATE LEVELS IN SKELETAL MUSCLE AND INTERMEDIATE TO

LOW LEVELS IN PANCREAS, KIDNEY AND PLACENTA. ACCORDING REF.1 ALSO

FOUND IN BRAIN, COLON, LUNG, SMALL INTESTINE, STOMACH, TESTIS, IN

SALIVARY GLANDS, THYROID GLANDS AND PROSTATE. VERY LOW LEVELS IN

PERIPHERAL BLOOD LYMPHOCYTES, FETAL BRAIN AND FETAL LIVER.

CC DEVELOPMENTAL STAGE: STRONGLY UP-REGULATED IN DIFFERENTIATING

FETAL MUSCLE CELLS (IN VITRO).

CC DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VMFA DOMAIN. INTEGRINS

WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.

Db 758 ----DHGEXLDDGMPPTLRSVPPFWNGCNEDEHCVPLDVLARSPLTAMEYQORVLRKP 813
 Qy 782 -----SIFESFMSLOCLVGGPREFWVTYVNDGDSYRTQVTFPPDLDSYRKVSTILO 836
 Db 814 ADCCSAYTLSTFTTFIESTQRVAVAEATLENGENAYSTVNLISQSANLQF--ASLIQ 871
 Qy 837 NORSORSWELACASASTEVSGALKASTCSINHPFIFFENSEVTFNITPDVDSKASLGNKL 896
 Db 872 KEDSDGS--IECVNEER-----RLQKQVCNVSPYPPFRAKAKVAFRLDFEF-SKSFILHHL 923
 Qy 897 LKANVTSENMPRTNKE--FQELPURYAVTWVTSHGVSTKY---LNFAS--ENTS 949
 Db 924 ETELAAAGSDNERDSTKEDNAPLRFHLKYEADVLFTRSSLSHSEYKLVNLSLEYDYG 983
 Qy 950 RVQHQYQVSNLGG--RSLPLSLVLPVRLNQTWIDRPVTPSEMLSSFTC-----HTK 1002
 Db 984 PFCSCIFRIQNLGLPFIHGMKTIPTATRSNRLKLRDLPT-DEVANTSCNIGWST 1042
 Qy 1003 ERLPSHDFLABRKAPVNVCSIAVCQRIQCIPFGIOEBPNATLKNLSFDWY----- 1057
 Db 1043 EYRPTFVE--EDLRAPQLNHSNDVSTNCNIRLVP-NQEIFHLLGNL---WLSLKA 1096
 Qy 1058 IKTSHNHLIVGTAEILFENDSVFTLLPGQGFVRSQRTKVEPPEVN-----PLPLIV 1111
 Db 1097 LYKSKMKNWALQGRFH-SPP-----IFREDPSRQIVFISQEDWQVPIWIV 1147
 Qy 1112 GSGVGLLILALITAAKYLGKFFK--ROYKMMSEGGPPGAP 1152
 Db 1148 GSTLGLLILALLVLAWKLGFFRSARRR-----PGLDP 1183

RESULT 11

IT2A2_BOVIN
 ID IT2A2_BOVIN STANDARD; PRT; 1170 AA.
 AC P53710;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Integrin alpha-2 precursor (Platelet membrane glycoprotein Ia) (GP1a)
 DE (Collagen receptor) (VLA-2 alpha chain) (CD49b) (Fragment).
 GN IT2A2.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OC NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94193647; PubMed=7511592;
 RA Kamata T., Puzon W., Takada Y.;
 RT "Identification of putative ligand binding sites within I domain of
 integrin alpha 2 beta 1 (VLA-2, CD49b/CD29).";
 RL J. Biol. Chem. 269:9659-9663(1994).
 CC !- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A RECEPTOR FOR LAMININ,
 CC COLLAGEN, COLLAGEN C-PROPEPTIDES, FIBRONECTIN AND E-CADHERIN. IT
 CC RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-P-G-E-R IN
 CC COLLAGEN. IT IS RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER
 CC CELLS TO COLLAGENS, MODULATION OF COLLAGEN AND COLLAGENASE GENE
 CC EXPRESSION, FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED
 CC EXTRACELLULAR MATRIX.
 CC !- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-2
 CC ASSOCIATES WITH BETA-1.
 CC !- SUBCELLULAR LOCATION: Type I membrane protein.
 CC !- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
 CC !- SIMILARITY: Belongs to the integrin alpha chain family.
 CC !- SIMILARITY: Contains 1 VWFA domain.
 CC !- SIMILARITY: Contains 7 FG-GAP repeats.

This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

EMBL; L25886; AAB59255.1; --
 DR PIR; I45914; I45914.
 DR HSSP; P17301; IACX.
 DR InterPro; IPR000413; Integrin_alpha.
 DR InterPro; IPR002035; VWF_A.
 DR Pfam; PF01839; FG-GAP; 3.
 DR Pfam; PF00357; integrin_A; 1.
 DR Pfam; PF00092; vwa; 1.
 DR SMART; SM00191; Int_alpha; 5.
 DR SMART; SM00327; VWA; 1.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 DR PROSITE; PS0234; VWFA; 1.
 KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 KW Platelet; Signal; Repeat; Polymorphism; Calcium; Magnesium.
 FT NON_TER 1
 FT SIGNAL <1 18 POTENTIAL.
 FT CHAIN 19 1170 INTEGRIN ALPHA-2.
 FT DOMAIN 19 1121 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 1122 1143 POTENTIAL.
 FT DOMAIN 1144 1170 CYTOPLASMIC (POTENTIAL).
 FT REPEAT 34 92 FG-GAP 1.
 FT REPEAT ? ? FG-GAP 2.
 FT DOMAIN 177 367 VWFA.
 FT REPEAT ? ? FG-GAP 3.
 FT REPEAT 423 475 FG-GAP 4.
 FT REPEAT 477 538 FG-GAP 5.
 FT REPEAT 540 599 FG-GAP 6.
 FT REPEAT 604 656 FG-GAP 7.
 FT CA_BIND 488 496 POTENTIAL.
 FT CA_BIND 552 560 POTENTIAL.
 FT SITE 472 474 CELL ATTACHMENT SITE (POTENTIAL).
 FT SITE 1146 1150 GFFKR MOTIF.
 FT DISULFID 72 81 BY SIMILARITY.
 FT DISULFID 669 726 BY SIMILARITY.
 FT DISULFID 778 784 BY SIMILARITY.
 FT DISULFID 854 865 BY SIMILARITY.
 FT DISULFID 1008 1039 BY SIMILARITY.
 FT DISULFID 1044 1049 BY SIMILARITY.
 FT CARBOHYD 94 94 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 101 101 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 332 332 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 421 421 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 449 449 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 464 464 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 688 688 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 748 748 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 945 945 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1063 1063 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1070 1070 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT VARIANT 580 580 G -> V.
 FT VARIANT 588 588 R -> K.
 FT VARIANT 725 725 R -> S.
 SQ SEQUENCE 1170 AA; 128929 MW; ECEFC5F2449FB1 CRC64;

Query Match 18.5%; Score 1099; DB 1; Length 1170;

Best Local Similarity 27.8%; Pred. No. 5e-65;
 Matches 340; Conservative 215; Mismatches 498; Indels 168; Gaps 47;

Qy 11 LTLCHGNLDTENAMTFQ-ENARFGQSVVQL----QGSRVVVGAPQISVAANQSGSLYQC 66
 Db 13 LNCVVAYVGLPKAKIFSGPSSEQGFVAVQGFIPKGNWLLVGSFWSGFPKRNVDVYKC 72
 Qy 67 --DYSTGSCPEIRIQ-----VPEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNT 117
 Db 73 PVDLSITTCCKLNLTQSTSMNSVTMTKMTNLSGLTLTRNVGTGGFLTCGPLWAOCCSQY 132
 Qy 118 YVKGCLCFGLFSGNLQQQPKFPEALRGCPQBDSDIAFLIDGSGSIIIPHDFRMKMFVSTVM 177

RT possible collagen-binding domain.";
RL J. Cell Biol. 109:397-407 (1989).
RN [2]
RP SEQUENCE FROM N.A.
RA Rieder M.J., Armel T.Z., Carrington D.P., Ozuna M., Kuldane S.A.,
RE Rajkumar N., Toth E.J., Yi Q., Nickerson D.A.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 168-368.
RA MEDLINE=98019223; PubMed=9353312;
RX Emley J., King S.L., Bergelson J.M., Liddington R.C.;
RT "Crystal structure of the I domain from integrin alpha2beta1.";
RL J. Biol. Chem. 272:28512-28517 (1997).
RN [4]
RP VARIANT HPA-5 (BR).
RX MEDLINE=94043762; PubMed=7901236;
RA Santos S., Kalb R., Walka M., Kiesel V., Mueller-Eckhardt C.,
RE Newman P.J.;
RL "The human platelet alloantigens Br(a) and Br(b) are associated with a
RT single amino acid polymorphism on glycoprotein Ia (integrin subunit
RL alpha 2).";
RN J. Clin. Invest. 92:2427-2432 (1993).
RN [5]
RP VARIANT GLU-534.
RX MEDLINE=20206009; PubMed=1074142;
RA Kroll H., Gardemann A., Fechter A., Haberbosch W., Santos S.;
RE "The impact of the glycoprotein Ia collagen receptor subunit A1648G
RT gene polymorphism on coronary artery disease and acute myocardial
RL infarction.";
RN Thromb. Haemost. 83:392-396 (2000).
CC -!- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A RECEPTOR FOR LAMININ,
CC COLLAGEN, COLLAGEN C-PROPEPTIDES, FIBRONECTIN AND E-CADHERIN. IT
CC RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-F-P-G-E-R IN
CC COLLAGEN. IT IS RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER
CC CELLS TO COLLAGENS. MODULATION OF COLLAGEN AND COLLAGENASE GENE
CC EXPRESSION, FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED
CC EXTRACELLULAR MATRIX.
CC -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-2
CC associates with beta-1. Interacts with HP55.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- POLYMORPHISM: Position 534 is associated with platelet-specific
CC alloantigen HPA-5 (BR). HPA-5A/BR(A) has Lys-534 and HPA-5B/BR(B)
CC has Glu-534. HPA-5B is involved in neonatal alloimmune
CC thrombocytopenia (NAIT or NAITP). The K534E polymorphism may play a
CC role in coronary artery disease (CAD).
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -!- DATABASE: NAME=PROV; NOTE=CD guide CD49b entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd49b.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: X17033; CAA34894.1; -
CC EMBL: AF512556; AAM34795.1; -
CC PIR: A33998; A33998.
CC PDB: 1A0X; 25-NOV-98.
CC PDB: 1DZ1; 02-AUG-01.
CC Genew: HGNC:6137; ITGA2.
CC MIM: 192974; -
CC GO: GO:0008305; C:integrin complex; TAS.
CC GO: GO:0005886; C:plasma membrane; TAS.
CC GO: GO:0004895; F:cell adhesion receptor activity; TAS.
CC GO: GO:0005518; F:collagen binding; TAS.
CC GO: GO:0007596; P:blood coagulation; TAS.
GO: GO:0007160; P:cell-matrix adhesion; TAS.
GO: GO:0007397; P:histogenesis and organogenesis; TAS.
interPro: IPR000413; Intergrin_alpha.
DR InterPro: IPR002035; VWFA.
DR Pfam: PF01839; FG-GAP 3.
DR Pfam: PF00357; Integrin_A; 1.
DR Pfam: PF00092; vwa; 1.
DR SMART: SM00191; Int_alpha; 5.
DR SMART: SM00327; VWFA; 1.
DR PROSITE: PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE: PS00234; VWFA; 1.
DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
RW Platelet; Signal; Repeat; Calcium; Magnesium; Polymorphism;
RW 3D-structure.
FT CHAIN 1 29
FT SIGNAL 30 1181
FT DOMAIN 30 1132
FT TRANSMEM 1133 1154
FT DOMAIN 1155 1181
FT DOMAIN 1155 1161
FT REPEAT 45 103
FT REPEAT ?
FT DOMAIN 188 378
FT REPEAT 378 433
FT REPEAT 434 486
FT REPEAT 488 549
FT REPEAT 551 610
FT REPEAT 615 667
FT CA BIND 499 507
FT CA BIND 563 571
FT CA BIND 627 635
FT SITE 1157 1161
FT DISULFID 83 92
FT DISULFID 680 737
FT DISULFID 789 795
FT DISULFID 865 876
FT DISULFID 1019 1050
FT DISULFID 1055 1060
FT CARBOHYD 105 105
FT CARBOHYD 112 112
FT CARBOHYD 343 343
FT CARBOHYD 432 432
FT CARBOHYD 460 460
FT CARBOHYD 475 475
FT CARBOHYD 699 699
FT CARBOHYD 1057 1057
FT CARBOHYD 1074 1074
FT CARBOHYD 1081 1081
FT VARIANT 534
FT TURN 170 171
FT STRAND 173 180
FT TURN 183 184
FT HELIX 188 199
FT TURN 200 201
FT STRAND 204 204
FT TURN 206 207
FT STRAND 209 216
FT STRAND 220 224
FT TURN 226 228
FT HELIX 232 240
FT TURN 241 241
FT HELIX 252 262
FT TURN 263 264
FT HELIX 266 268
FT TURN 269 269
FT STRAND 275 282
FT HELIX 289 291
FT HELIX 292 301
FT TURN 302 303
FT STRAND 304 311
FT HELIX 313 317

```
FT TURN 318 319
FT HELIX 323 330
FT TURN 331 332
FT HELIX 337 340
FT STRAND 341 344
FT HELIX 347 353
FT HELIX 354 362
FT TURN 363 363
SQ SEQUENCE 1181 AA; 129295 MW; 7ELB7ED968A94070 CRC64;

Query Match 18.3%; Score 1087.5; DB 1; Length 1181;
Best Local Similarity 27.3%; Pred. No. 3e-64;
Matches 339; Conservative 214; Mismatches 498; Indels 193; Gaps 44;

QY 3 LRVLLLTALT-----LCHGNLNTENAMTFO-QNARGFGOSVQL---QGRVWVCAPOE 53
DB 11 LPLLLVLAUSQGLNCLLAATVGLPEAKLFGSPSSSEFGVAVQVQPINPKGNWLLVGSPTS 70
QY 54 IVAANQSGSLYQC--DYSTGSCBPIRLQ-----VPVEANMNSGLSLAATSPQOLLA 104
DB 71 GPPENEMGDYKCPVDLSTATCEKLNLTSTSPNVTENKNNLSGLILTRNNGTGGFLT 130
QY 105 CGPTVHTQCSNTYVYVGLCPFLGSLNLRQOQKPEALRGCPQEDSDIAFLDGSGLIHP 164
DB 131 CGPLWAOQGNQYTTGVCSDISDPF-QLSASFSPATQPCPSL-IDVVVVCDESSSIYEW 188
QY 165 DFRMKFVSTVMEQLK--KSKTLFSLMOYSEEFRIHFTKFPQNNPNRSLVKPTQLL 222
DB 189 D--AVXNFLEKFQGLDIGTKTQVGLIOYANNPRVFNLTNTYKTEEMIVATSQSYG 246
QY 223 G-RTHATGIRKVVRLFNITNARKNAKILVITDGEKFGDPLGYEDVPEADREGVI 281
DB 247 GDUINTFGALQYARKYAYSAAGRSRATKVMVVTGDSH-DGSMKAKVIDQCNDHNL 305
QY 282 RYVIGV-----GDAPFSEKSRBLNTIASKPRDRHVQVNNFEALKTIONLREKIPAI 336
DB 306 RFGIAGVLYNRLNALTOKNLKEIKALASIPTEPYFNWSDRAALKAGTLGEQIFSI 365
QY 337 GTGTGSSSFEHMSQEGSAAITSNP--LLTSGSYWAGSVFLYTSKSTPTINMT- 393
DB 366 GTVQGG-GDNFMQMSQVGSADYSQSDIIMLGAVGAFGMSGTIVQKTSHGHLFPKQAF 424
QY 394 -RVSDMN-DAYLGAAATILNRVQSLVGLAPYQHIGLVAMFRONTGWESNANV--- 448
DB 425 DQILODRNHSYLGYSVAALSTGESHFVAGAPRANTQIVLYSN-----ENGNITVI 479
QY 449 ---KGTQIGAYFGASICSVDVDSNGSTDLVIGAPHYEQTR--GGQVSVCPPLPQORAR 503
DB 480 QAHRGDIQSGYFSGVLCSDVDVDKDTITDVLVAGAPMYSDLKKEGRVYLFYIKKGLQ 539
QY 504 WQCDVLYGSGQOPWGRFGAALTVLGVNGDKLTDVAIGAPCEEDNRGAVYLPHTSGSG 563
DB 540 HQ---FLEGPEGLTFRFSAIALSDINWDGNDVIVGSPLENQSGAVIYNGHQT- 595
QY 564 IGFHSQRIAGS--KLSPLQYFGQSLGGQDLTMDGLVLTGAGQHVLLRSQPVLV 621
DB 596 IRTKYKILGSDGAFRSHLYQFGRSLDGVGLNGDSITDVSIGAPGVVQLWSQSIADV 655
QY 622 KAIMENPREVARNVECHDQVYKGEARVCLHVQKSTEDRLBGGQIQSVVYDLAL 681
DB 656 ALEASTPPEKI--TLVNKNAQII-----LKLCF---SAKPRPTKQNNQAVIYVNTL 702
QY 682 D-----SGRPHSRVAFVETKNSRTRQTVLGLTQTC--ETLKLQPLNCIEDPSPVILRLN 735
DB 703 DADGPFSSRVTSRGLFKENNERCLQKNMVVYVQAGSCEPHIYIQEPS---DVVNSLDLRVD 759
QY 736 FSLVGTPLSFAFNLRLVLAEDQRLFTALFPFEKNCNDNIQDDLSITF-----SFMSLD 791
DB 760 ILENPGTS-----PALEAYSETAKVPSIPFKDGGEDGLCISDLVDVROIQAQEQ 813
QY 792 CLVVGCPREFNVTVVRNDEGDSYRQVTFEFTPLDLSYRKVSTLQNRQSRWRLACRSA 851
DB 814 FIVSNQKRLTSTVTLKNEBSAYNTGIIVVDFSENLF-----ASFSLPVDGT 861
```


CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

```

CC -----
CC ENBL; AF074015; AAC31952.1; .
CC DR EMBL; AF112345; AAF21944.1; .
CC DR EMBL; AF172723; AAF61638.1; .
CC DR HSP; P17301; LAOX.
CC DR Genew; HGNC:6135; ITGA10.
CC MLM; 604042; .
CC DR GO; GO:0008305; C:integrin complex; TAS.
CC DR GO; GO:0004895; F:cell adhesion receptor activity; TAS.
CC DR GO; GO:000518; F:collagen binding; TAS.
CC DR GO; GO:0007160; P:cell-matrix adhesion; TAS.
CC DR InterPro; IPR000413; Integrin_alpha.
CC DR InterPro; IPR002035; WWP_A.
CC DR Pfam; PF01839; FG-GAP; 3.
CC DR Pfam; PF00092; vwa; 1.
CC DR PRINTS; PR01185; INTEGRINA.
CC DR SMART; SM00453; VWFADOMAIN.
CC DR SMART; SM00191; Int_alpha; 4.
CC DR SMART; SM00327; WVA; 1.
CC DR PROSITE; PS00242; INTEGRIN_ALPHA; FALSE_NEG.
CC DR PROSITE; PS00234; WVFA; 1.
CC KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Signal; Repeat; Calcium; Magnesium.
FT SIGNAL 1 22
FT CHAIN 23 1167
FT DOMAIN 23 1122
FT TRANSMEM 1123 1145
FT DOMAIN 1146 1167
FT REPEAT 38 97
FT REPEAT 167 350
FT DOMAIN 167 350
FT REPEAT 365 427
FT REPEAT 428 482
FT REPEAT 483 545
FT REPEAT 546 605
FT REPEAT 608 660
FT DOMAIN 1134 1140
FT CA_BIND 494 502
FT CA_BIND 558 566
FT CA_BIND 620 628
FT DISULFID 76 86
FT DISULFID 666 675
FT DISULFID 681 736
FT CARBOHYD 98 98
FT CARBOHYD 234 234
FT CARBOHYD 336 336
FT CARBOHYD 364 364
FT CARBOHYD 733 733
FT CARBOHYD 763 763
FT CARBOHYD 839 839
FT CARBOHYD 921 921
FT CARBOHYD 1011 1011
FT CARBOHYD 1018 1018
FT CARBOHYD 1039 1039
FT CONFLICT 844 844
FT CONFLICT 909 909
FT CONFLICT 926 926
SQ SEQUENCE 1167 AA; 127573 MW; AE7D3A1C25C1AEAO CRC64;
Query Match 18.2%; Score 1085.5; DB 1; Length 1167;
Best Local Similarity 29.6%; Pred. No. 3.9e-64;
Matches 366; Conservative 196; Mismatches 492; Indels 183; Gaps 47;
QY 3 LRVLITLALTLCHGFLNDTNAMTFQENARG-FQGVVQLQGR---VVVGAPQEIIVAAV 58
DB 11 LPLVLTG--LCSPFLNDEHHPRLFGPPEAEFGYSVLQHVGGQQRWMLVGAPWDGPGSD 68
QY 59 QRGLSYOC-----DYSTG-SCEPIRLQVPVEAVNMSLGLSLAATSPQ 101

```

```

Db 69 REGDVYRCPVGAHNAPCAKHLGDYQLGNSHP-----AVNHLGMSILETDGSG 120
QY 102 LLAQCPVTHQTCSENTYKGLCFPGSNLRQOPKPEALRCCPQSDSDIAFLIIGSSSI 161
Db 121 FWACAPLWSRACGSGVFSSGICARVDASQFGSLAPTAQR-CPTY-MDWVIVLVDGNSI 178
QY 162 IPHD-----FRM--KEFVSTVMEQLKSKTLFSLMOYSEEFRIHFTKFKFQNNPNSRL 214
Db 179 YWSEVQTFPLRLVCKLFIDP--EQIQ-----VGLVQGESPVHMSLGDFTKEEVVRA 231
QY 215 VPIITQLLGR-THATGIRKVVRELFNITNGARKNAFKILVITDGEKF-GDPLGYEDVI 272
Db 232 ANLSRREGRETKAQAINVACTEGSQSGHGRPEARLLVVVTDGSDGSELPAALKA 291
QY 273 PADREGVIRVIGV-GDAFSEKS-----ROSLNTIASKPRDHVFOVNNPALKTIQNG 327
Db 292 CEAGR-VTRGIAVLGHYLRQRDPSSFLREIRTIASDPERFFNTVDEAALTDVDA 349
QY 328 LREKIFAJEGTQSGSSFEHMSQEGFSAATISNGPLLSVTVSGVDWAGGVLYTSKKS 387
Db 350 LGDRIFGLEGSHAENESSFGLMSQIGFSTRLKDGILFGMVGYADWGSVLWLSGGHRL 409
QY 388 TFINTRVDS-----DMNDAYLGYA-AAIILNRVQSLVLAGAPRYCHILGLVAMPR-ONTG 440
Db 410 FPRVVALEDEFPALONHAAYLGISVSMILRGERRFLSGAPRFRHKGKVIAPOLKKG 469
QY 441 MWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYY--EOTRGGOVSVCPUPR 498
Db 470 AVRVAQSLQGEQIGSYFGSELPLDTRDGTDLVLAAPMPGLPQNKETGRVYVLV-- 527
QY 499 GQARWQCDVLYGQGGQWGRFGRGALTVLDGVNGDKLTDVAIGAPGEDNRGAVYLFHG 558
Db 528 GQOGLTLQGLTQPEPPD-ARFGFAMGALPDLDODGADVAAGAPLEDHGOGALYLYHG 586
QY 559 TSGGISPSHSORIASGKSLSPRLQYFGQSLSGQDLTMDGLVLTGVAQGHVLLRSOPV 618
Db 587 TQ-SGVREHPAQRIAAASMPHALSYFGESVDGLDGLDGLDVAVAGAGAAIILSRPI 645
QY 619 LRVKALMEFNPREVARNVFCNDQVVKGEAG--EVRVCLHVQKSTRDLRSGQSQSVYT 676
Db 646 VHLTFSLEVTPOAISVQVORDCR--RQGEAVCLTAALCFQVTSRTPGWDH---QFYNR 699
QY 677 YDLALDSGRPHSRVAFNET--KNSTRRTQVLGLTQTCETLKLQLPNCIEDVPSVILRL 734
Db 700 PTASLDEWTAGARAADFQSGQRLSPRLRLSVG-NVTCQLHFLVLD-TSDYLRPVALTV 757
QY 735 NFSLVGTPLSPAGNLRPVLAEDAQLFTALPPEKXCGDNICQDLSITPFSMLDC--- 792
Db 758 TFLDNTTKPG-----PVLNBSGPTSIQKLVPFSCDGPNECVTLVLQ-----VMDIRG 809
QY 793 -----LVVGGPREFNVTVVRNDCGDSYRTQVTFPFLDLISYRKVSTLQNSQSRWL 846
Db 810 SRKAPFVVRGGRKVLVSTTLNKENAYNTSLIIFSRNL---HLASITPQR-BSPIKV 865
QY 847 ACESASSTEVSGALKSTCSINHPITFPENSEVTNITFDVDSKASLG---NKLL-----L 898
Db 866 ECAAPSA-----HARLCSVGHVPVQCAKVTFLLEFEFFSCSLLSQVFGKLTASSDSL 918
QY 899 KANVTSENNMPTNKTETQLELPVKYAVVMVTVSHGVSTKYLNFTASENTSRVMOHOY-- 956
Db 919 ERNGTLQENTAGT-----SAYIQEYH-----LLFSESTLHRYEYVHPYGT 959
QY 957 -----QVSNLG---QRSPLPISLVPLVP-----VRINQVTIWDPRQVTF 991
Db 960 LPVGPGEFKTTLRVQLNGCYVVSGLIISA--LLPAVAHGNGYFLISQVI----- 1008
QY 992 SENLSTCHTKERLPSHSDFLAELKAPVWNCISVACORIQCDIPFGQIEBENATLKN 1051
Db 1009 TNNAACIVONLTPPGPPVHPPEELQHTRLNGSNTOCQVVRCHLGOLAKGTSEVGLLRL 1068
QY 1052 LSFDFWTKSHHMLIVSTABILFVDSVFTLLFGQAGFVRSQETKVPPEPVNPLPLTV 1111
Db 1069 VENEFFERRAKFKSLTVVSTFELGTGEEGVLQLTASRWSESLELV-VQTRPILISILILI 1127

```


Db 311 GHYNRGNLSTKVFBEIKSIASRPTKHPFNVSDELALVTIVKALGERIFALEATADQSA 370
Qy 344 SSFEHEMSQGFSAATSNOPLLSTVGSYDMAGVFLYTSKEKSTFINMT--RVDSMDND 401
Db 371 ASFEHEMSQGFSAHYSQDMVLMGAVGAYDNGTVVWQXANQWVPHNTFFQTEPAKXNE 430
Qy 402 ---AYLGYAAAIILNRVQSLVLCAPRYQHIGLVAMFRQNTGMWESNANVKGTQIGAYFG 458
Db 431 PLASYLGYTVNSATIPCDVLYIAGQRYNHTGQVVIYKMGEDGNINILQTLGGEQIGSYFG 490
Qy 459 ASLCSVDVDSNGSTDLVIGAPHY-----YEQTR-GQVSVCPPLPRGQRA 502
Db 491 SVLTIDIDKSDYTDLLLVGAPMYMGTEKEEQKVYVYAVNTRPYQMSLEPIROTCCS 550
Qy 503 RWOCDAVLYGEOGQPMG-RFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLPHTGSG 561
Db 551 SLKDNSTCKENKNEPCGARGETAIAVKDLNVGDFNDVIGAPLEDDHAGAVYIYHG-SG 609
Qy 562 SGISPSHSQRIAGSKUSPRIQYFGQSLSGQDQTLMDGLVDLTVGAQGHVLLLRSPVLRV 621
Db 610 KTIREAYAQRIPSGGDGKTLKFFQSGIHGEMDLNGDGLTDVTIGGLGAALFWARDVAVV 669
Qy 622 KATMEENPREVARNVPCNDQVVKGEAG--EVRVCLHVO-KSTRDLREGQIQSVVTYD 678
Db 670 KVTWNEFPKVNIOKQNCR---VEGKETVCINATKCFHVXKLKSKESIIYADLQ----YR 722
Qy 679 LALDSGRPHSRAVFNET-----KNSTRQTVLGLTQTCETLKLQLPNCI----- 723
Db 723 VTLDSLQIISRSFSGTOERKIOQNIIVRESE-----CIRHSFYMLQK 765
Qy 724 EDVPSPIVBLNLSVGLTSLAFGNLSPVLAEDAQRLLPTALPPEKXGNDNICQDDLSI 783
Db 766 HDPQDSVRVTLDFNLT-DPENG-----PVJODALPNSVHEHIIPFAXDCGKXKICISDLTL 819
Qy 784 TFSFMSLCLVVGQPRE-FNVTVTRNDGBDSYNTQVTFPPPLDLSYRKVSTLQNRQSR 842
Db 820 NVSTTEKSLIVKQCHDKFNVLTVXNKGDSAYNTRVWQHSNLIPISGIEBIOKQ---- 875
Qy 843 SWRLACESASTEVSGALKSTCSINHPFIPENSEVTFNITFDVDSKASLGKLL-LKAN 901
Db 876 ----SCSEN-----QNITCRVGYFFLRAGETVTFKIIFOPTSHLSENAIHLSAT 922
Qy 902 VTSENNMPTNKTEFQLELPVKYAV---YMWVTSHGVS-----KYLNFASENTSR 950
Db 923 SDSEEPLESLNDNEVNISSIPVKEYEVLQFYSSASEHHISVAANETIPEFINST--EDIGN 980
Qy 951 VMQHQYQVSNLQGRSLP---ISLVP-----LVPVRLNQTVIWD-----RP--- 987
Db 981 EINVPYTIKRGHPMPPELQLSISFPNLTDGYPVLYPIG-----WSSSDNVNCRPSL 1034
Qy 988 -----QVTFPS-----ENLSSTCHTKERLPSHSDFLAELRKAPVVCNSIAV 1027
Db 1035 EDPFGINSKXOTISKSEVLKRGTIQCSSTC-----GVATITCSLLP 1077
Qy 1028 CQRIQCDI-----PFFGIOBEF---NATLKGNSLFDWYKTSNNHLLIVSTABILEFDS 1078
Db 1078 SDSLQVNVVLLWKPTF-IRAHFSSNLMLTLAGELK-----SENSSLTSSN----- 1123
Qy 1079 VFTLLPGQAVRQTEKVPPEVNPPLP-IVGSSVGLLALLALITAAALYKLGPFKR 1136
Db 1124 -----RKRELAIQISKDGLPGRVPLWVILLSAFAGLLMLLILALWKIGPFKR 1172
Qy 1137 QYKDMSE 1144
Db 1173 PLKKMEK 1180

Search completed: June 7, 2004, 17:12:38
Job time : 18.1123 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 16:56:30 ; Search time 40.0446 Seconds
(without alignments)

9084.693 Million cell updates/sec

Title: US-09-902-481b-1

Perfect score: 5956

Sequence: 1 MALRVLLTALTLCHGFNLD.....FKRQYKDMWSEGGPPGABPQ 1153

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL 25.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4459	74.9	1151	11 Q9JI30	Q9JI30 rattus norv
2	3976.5	66.8	1036	11 Q8CA73	Q8CA73 mus musculus
3	3824.5	64.2	920	6 Q28984	Q28984 sus scrofa
4	3514	59.0	1169	4 Q8IVA6	Q8IVA6 homo sapien
5	3352.5	56.3	1169	11 Q9QXH4	Q9QXH4 mus musculus
6	3282.5	55.1	1161	11 Q9QYE7	Q9QYE7 rattus norv
7	1545	25.9	1161	11 Q9WTV4	Q9WTV4 mus musculus
8	1534.5	25.8	1160	11 Q8R200	Q8R200 mus musculus
9	1419	23.8	1196	13 Q8RTP1	Q8RTP1 cyprinus ca
10	1368.5	23.0	1187	13 Q8RTP0	Q8RTP0 cyprinus ca
11	1361	22.9	1086	4 Q9GHB1	Q9GHB1 homo sapien
12	1277	21.4	927	6 Q8HZV0	Q8HZV0 bos taurus
13	1171.5	19.7	1167	11 Q88340	Q88340 rattus norv
14	1125.5	18.9	1167	11 Q88341	Q88341 rattus norv
15	1068.5	17.9	1171	13 Q42094	Q42094 gallus gall
16	1060.5	17.8	1168	11 Q7IQC3	Q7IQC3 mus musculus

17	1049	17.6	1038	11 Q8BS01	Q8BS01 mus musculus
18	1029.5	17.3	1160	6 Q8MKF4	Q8MKF4 felis silve
19	1020	17.1	895	11 Q9WUF8	Q9WUF8 mus sp. itg
20	865	14.5	348	4 Q8TES5	Q8TES5 homo sapien
21	852	14.3	1332	5 Q9BPQ8	Q9BPQ8 halocynthia
22	816	13.7	205	11 Q63001	Q63001 rattus norv
23	753.5	12.7	780	13 Q66271	Q66271 xenopus lae
24	738	12.4	823	4 Q8W118	Q8W118 homo sapien
25	686.5	11.5	823	11 Q8CE84	Q8CE84 mus musculus
26	683	11.5	1032	11 Q61989	Q61989 mus musculus
27	651.5	10.9	257	11 Q8C270	Q8C270 mus musculus
28	643	10.8	1036	11 Q91YD5	Q91YD5 mus musculus
29	641.5	10.8	1033	6 Q95GU3	Q95GU3 bos taurus
30	637	10.7	1474	5 Q86G87	Q86G87 pseudoplusi
31	619.5	10.4	1041	5 Q9UB90	Q9UB90 lytechinus
32	614.5	10.3	1041	5 Q76378	Q76378 lytechinus
33	589	9.9	1034	13 Q98T77	Q98T77 gallus gall
34	579.5	9.7	1054	5 Q9U6S1	Q9U6S1 strongyloce
35	555.5	9.3	1053	11 Q80YP5	Q80YP5 mus musculus
36	550	9.2	1033	13 Q42598	Q42598 xenopus lae
37	546	9.2	1036	6 Q7YRP8	Q7YRP8 equus cabal
38	535.5	9.0	1016	13 Q91779	Q91779 xenopus lae
39	530	8.9	974	11 Q924W2	Q924W2 rattus norv
40	529	8.9	1073	11 Q8CC06	Q8CC06 mus musculus
41	529	8.9	1119	5 Q86G88	Q86G88 pseudoplusi
42	526	8.8	1047	6 Q9WZD6	Q9WZD6 bos taurus
43	525.5	8.8	1007	6 Q9GK48	Q9GK48 bos taurus
44	522.5	8.8	1132	11 Q80Z18	Q80Z18 mus musculus
45	514.5	8.6	1034	6 Q9TUM4	Q9TUM4 oryctolagus

ALIGNMENTS

RESULT 1

Q9JI30 PRELIMINARY; PRT; 1151 AA.
AC Q9UI30
DF 01-OCT-2000 (TREMREL. 15, Created)
DT 01-OCT-2000 (TREMREL. 15, Last sequence update)
DE Integrin beta 2 alpha subunit.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Fathallah D.M. St., Zeria K. Jr.;
RT Cloning of the rat CD11b cDNA sequence."
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF268593; AAF81280.1; -
DR HSP; P11215; 1BHQ.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; P:cell adhesion receptor activity; IEA.
DR GO; GO:0007140; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VMP_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS50234; VFWA; 1.
SQ SEQUENCE 1151 AA; 126943 MW; 8P785695D4074CA5 CRC64;

Query March 74.9%; Score 4459; DB 11; Length 1151;

Best Local Similarity 73.5%; Pred. No. 2.9e-315;

Matches 847; Conservative 144; Mismatches 160; Indels 2; Gaps 2;

QY 1 MALRVLLTALTALCHGFNLDNTENAMTFQENARGFGQSVVQLQSGRVVVGAPQEIIVANQR 60
DB 1 MTLKALEYVLTALCHGFNLDNTENAMTFQENARGFGQSVVQLQSGRVVVGAPQEIIVANQR 60
QY 61 GSLYQCDYSTGSCBPIRQVPEAVNMSLGLSLAATSPPOLLAGCPVTHQTCSENTYVK 120
DB 61 GALYQCDYSTGSCBPIRQVPEAVNMSLGLSLAATSPPOLLAGCPVTHQTCSENTYVK 120
QY 121 GLCFLFGSNLRQOQKPPPEALRGCPQEDSDIAPLIDGSGSIIPHDPRMKFVSTVMEQL 180
DB 121 GLCYLFGSNLRQOQKPPPEALRGCPQEDSDIAPLIDGSGSIIPHDPRMKFVSTVMEQL 180
QY 181 KKSXTLFSLMQYSEBERTHFTFKFQNNPNSLVKPTTOLLGRTHATGIRKVVRELFN 240
DB 181 KKSXTLFSLMQYSEBERTHFTFKFQNNPNSLVKPTTOLLGRTHATGIRKVVRELFN 240
QY 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVIPADREGVIRYVIGVDAPRSEKSRQEL 300
DB 241 KINGARDNAKILVITDGEKFGDPLGYEDVIPADREGVIRYVIGVDAPRSEKSRQEL 300
QY 301 NTIASKPRDHVQVNNFEALKTIQNLREKIPAIETGTGSGSSPHEHMSQGFSAIT 360
DB 301 DTIASKPRDHVQVNNFEALKTIQNLREKIPAIETGTGSGSSPHEHMSQGFSAIT 360
QY 361 SNGPILLSVGSYDWDAGVFLYTSKXSTFINMTVRVDSMDAYLGYAAAIILNRVQSLV 420
DB 361 SNGPILLSVGSYDWDAGVFLYTSKXSTFINMTVRVDSMDAYLGYAAAIILNRVQSLV 420
QY 421 LGAPRYQHIGLVAFRONTGWESNANVKQTQIGAYFGASICSVDVDSNGSTDLVIGAP 480
DB 421 LGAPRYQHIGLVAFRONTGWESNANVKQTQIGAYFGASICSVDVDSNGSTDLVIGAP 480
QY 481 HYETQIRGGVSCPLPRGORAKWQCDVLYGEGQPGWGFAGALTIVLGDVNGDKLTDVA 540
DB 481 HYETQIRGGVSCPLPRGORAKWQCDVLYGEGQPGWGFAGALTIVLGDVNGDKLTDVA 540
QY 541 IGAPBEDNRGAVLHGTGSGSIPSHSRIAGSLSPRLQYFGOSLQSGQDLTMDGLV 600
DB 541 IGAPBEDNRGAVLHGTGSGSIPSHSRIAGSLSPRLQYFGOSLQSGQDLTMDGLV 600
QY 601 DLTVGAQGHVLLARSQVLRVKAIMEFNPREVARNVPECNDOVVVKEAGEVRCVLRVOK 660
DB 601 DLTVGAQGHVLLARSQVLRVKAIMEFNPREVARNVPECNDOVVVKEAGEVRCVLRVOK 660
QY 661 STDRLEGOIQSVTYVDLALDSGRPHSAVFNETKNSRTRQVLTGTCETLKLQLP 720
DB 661 STDRLEGOIQSVTYVDLALDSGRPHSAVFNETKNSRTRQVLTGTCETLKLQLP 720
QY 721 NCIEDPVSPIVLRNLSVGTPLSAPGNLRPLVLAQRLFTALPFEKNCNDTICODD 780
DB 721 NCIEDPVSPIVLRNLSVGTPLSAPGNLRPLVLAQRLFTALPFEKNCNDTICODD 780
QY 781 LSITFSFMSLCLVVGPGREFNVTVTRNDGDSYRTQVTFPPPLDLRYKRVSTLQNR 840
DB 781 LSITFSFMSLCLVVGPGREFNVTVTRNDGDSYRTQVTFPPPLDLRYKRVSTLQNR 840
QY 841 QRSWRLACESASTESVGLKSCSINRPIPENSEVTFNITFDVDSKASLGNKLLKA 900
DB 841 QRSWRLACESASTESVGLKSCSINRPIPENSEVTFNITFDVDSKASLGNKLLKA 900
QY 901 NVTSNNMPTNKTEFQLELPVYAVVMVTVSHGVSTKVLNFTASNTSRVMOHOVQSN 960
DB 901 NVTSNNMPTNKTEFQLELPVYAVVMVTVSHGVSTKVLNFTASNTSRVMOHOVQSN 960
QY 961 LGORSIPISLVLPVRLNQTIVDRPQVTFPSNLSSTCHTKERLPSHSDFLAELRKAPV 1020
DB 961 LGORSIPISLVLPVRLNQTIVDRPQVTFPSNLSSTCHTKERLPSHSDFLAELRKAPV 1020
QY 1021 VNCIAVCORIQCDIPFGIOBEFNATLKNLSFDWYIKTSHNLLIVSTAILFNDSVP 1080
DB 1021 VNCIAVCORIQCDIPFGIOBEFNATLKNLSFDWYIKTSHNLLIVSTAILFNDSVP 1080
QY 1081 TLLPGQGAFVRQSTETKVEFFEPVNPPLPLIVGSSVGLLALLALITAALYKLGFFKRYND 1140
DB 1081 TLLPGQGAFVRQSTETKVEFFEPVNPPLPLIVGSSVGLLALLALITAALYKLGFFKRYND 1140

DB 1079 ALLPQGBTEFVCAQTEKVEPYTVHNPVPLVGVSSGVLVLLALITAGLYKLGFFKRYND 1138
QY 1141 MNSEGPGGABRQ 1153
DB 1139 MNNEGQDGPQ 1151
RESULT 2
Q8CA73 PRELIMINARY; PRT; 1036 AA.
ID Q8CA73
AC Q8CA73
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Integrin alpha M.
GN ITGAM OR F730045J24RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CS7BL/6J; TISSUE=Spinal cord;
RX MEDLINE=22354683; PubMed=12466851;
RA The PANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573 (2002).
DR EMBL; AK039444; BAC30350.1; --
DR PIR; PT0572; PT0572.
DR PIR; PT0633; PT0633.
DR PIR; PT0697; PT0697.
DR MGD; MGI:96607; Itgam.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; P:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 1.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWFA; 1.
SQ SEQUENCE 1036 AA; 115126 MW; 4F4F39BF1E188E77 CRC64;
Query Match 66.8%; Score 3976.5; DB 11; Length 1036;
Best Local Similarity 66.8%; Pred. No. 3.3e-280;
Matches 771; Conservative 120; Mismatches 144; Indels 119; Gaps 2;
QY 1 MALRVLLTALTALCHGFNLDNTENAMTFQENARGFGQSVVQLQSGRVVVGAPQEIIVANQR 60
DB 1 MTLKALEYVLTALCHGFNLDNTENAMTFQENARGFGQSVVQLQSGRVVVGAPQEIIVANQR 60
QY 61 GSLYQCDYSTGSCBPIRQVPEAVNMSLGLSLAATSPPOLLAGCPVTHQTCSENTYVK 120
DB 61 GALYQCDYSTGSCBPIRQVPEAVNMSLGLSLAATSPPOLLAGCPVTHQTCSENTYVK 120
QY 121 GLCFLFGSNLRQOQKPPPEALRGCPQEDSDIAPLIDGSGSIIPHDPRMKFVSTVMEQL 180
DB 121 GLCYLFGSNLRQOQKPPPEALRGCPQEDSDIAPLIDGSGSIIPHDPRMKFVSTVMEQL 180
QY 181 KKSXTLFSLMQYSEBERTHFTFKFQNNPNSLVKPTTOLLGRTHATGIRKVVRELFN 240
DB 181 KKSXTLFSLMQYSEBERTHFTFKFQNNPNSLVKPTTOLLGRTHATGIRKVVRELFN 240
QY 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVIPADREGVIRYVIGVDAPRSEKSRQEL 300
DB 241 KINGARDNAKILVITDGEKFGDPLGYEDVIPADREGVIRYVIGVDAPRSEKSRQEL 300

```
QY 301 NTIASKPRDHPVQVNFZALKTIQNLREKIPAIETGTGTGSSSSPEHEMSQEGFSAIT 360
DB 301 DTIASKPAGHPVQVNFZALKTIQNLREKIPAIETGTGTGSSSSPEHEMSQEGFSAIT 360
QY 361 SNGPLSTVGSYDAGVFLYTSKSTFIMNTRVDSMDNDAYLGAAAIILNRVQSIV 420
DB 361 SNGPLSGVSGPDWAGAFLYTSKDVFTINTTRVDSMDNDAYLGAAVILNRVQSIV 420
QY 421 LGAPRYCHGLVAMFRQNTGWSNANVKTQIGAFGASLCSVDVDSNGSDTLVLIGAP 480
DB 421 LGAPRYCHGLVAMFRQNTGWSNANVKTQIGAFGASLCSVDVDSNGSDTLVLIGAP 480
QY 481 HYVEQTRGGQVSCPLPRGQARWQCDVLYGEGQPGWGRFGAALTVLGDVNGDKLTDVA 540
DB 481 HYVEQTRGGQVSCPLPRGQARWQCDVLYGEGQPGWGRFGAALTVLGDVNGDKLTDVA 540
QY 541 IGAPGEDNRGAVILFHGTGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMGHV 600
DB 541 IGAPGEDNRGAVILFHGTGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMGHV 600
QY 601 DLTVAQGHVLLRSOPVLVRKAIMFPNPREVARNVPECNDQVVGKEAGEVRVCLVQK 660
DB 601 DLTVAQGHVLLRSOPVLVRKAIMFPNPREVARNVPECNDQVVGKEAGEVRVCLVQK 660
QY 661 STRDLREGQIQSVVYTDLALDSGRPHSRVAFNETKSTRTQVGLGTQTCETLKLQLP 720
DB 661 STRDLREGQIQSVVYTDLALDSGRPHSRVAFNETKSTRTQVGLGTQTCETLKLQLP 720
QY 721 NCIEDPVSPIVLRNLSVGTPLSAPGNLPUVAEDAQALFALPFPEKNCNDNIQQD 780
DB 721 NCIEDPVSPIVLRNLSVGTPLSAPGNLPUVAEDAQALFALPFPEKNCNDNIQQD 780
QY 781 LSITFSMGLDCLVWGCPREFVTVVRNDGDSYRTQVTFPFLDLSYRKVSTLQNRS 840
DB 781 LSITFSMGLDCLVWGCPREFVTVVRNDGDSYRTQVTFPFLDLSYRKVSTLQNRS 840
QY 841 QRSWL-ACESASSTEVSGALSTCSINHPIPPENSEVTMTFVDSKASLGNKLLK 899
DB 841 QRSWL-ACESASSTEVSGALSTCSINHPIPPENSEVTMTFVDSKASLGNKLLK 899
QY 900 ANVTSNNPRTNKTKEFQLELPVYAVMVVTSVSHGVTYLNFTASENTSRVMOHQYQVS 959
DB 900 ANVTSNNPRTNKTKEFQLELPVYAVMVVTSVSHGVTYLNFTASENTSRVMOHQYQVS 959
QY 960 NLGQSLPISVFLVPLRNQTVIMDRPQVTFSENLSSTCHTKERLPSHDSFLAELKAP 1019
DB 960 NLGQSLPISVFLVPLRNQTVIMDRPQVTFSENLSSTCHTKERLPSHDSFLAELKAP 1019
QY 1020 VVNCSTAVCORIQCDIPFGIOEEFNATLKNLSFDWYIKTSHNLLIYSTAIIIFNDV 1079
DB 1020 VVNCSTAVCORIQCDIPFGIOEEFNATLKNLSFDWYIKTSHNLLIYSTAIIIFNDV 1079
QY 1080 FTLLPQAGFVRSQETETKVEPEVFNPLPLIYGVSSVGGILLALITAAALYKLGFFKQYK 1139
DB 1080 FTLLPQAGFVRSQETETKVEPEVFNPLPLIYGVSSVGGILLALITAAALYKLGFFKQYK 1139
QY 1140 DMSGEGGPPGAEPO 1153
DB 1140 DMSGEGGPPGAEPO 1153
QY 1023 DMWNEAAPQDAPPQ 1036
DB 1023 DMWNEAAPQDAPPQ 1036
```

RESULT 3

Q28984

ID Q28984 PRELIMINARY; PRT; 920 AA.

AC Q28984;

DT 01-NOV-1996 (TReMBLrel. 01, Created)

DT 01-FEB-1997 (TReMBLrel. 02, Last sequence update)

DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)

DE CD11b (Fragment).

GN CD11b.

OS Sus scrofa (Pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Lee J.-K., Schook L.B., Rutherford M.S.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U40072; AAB16869.1; -.
DR HSSP; F11215; 1BHO.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS50234; VWF_A; 1.
FT NON_TER 920
FT NON_TER 920
SQ SEQUENCE 920 AA; 102440 MW; E96CC51E350DD5AC CRC64;

Query Match 64.2%; Score 3824.5; DB 6; Length 920;
Best Local Similarity 79.8%; Pred. No. 3.2e-269;
Matches 735; Conservative 77; Mismatches 108; Indels 1; Gaps 1;

QY 134 POKPEALRGCPQSDSDIAPLIDGSGSIIPHDPRMKGFSVSTMQLKKSKTLPALMOYS 193
DB 1 POKPEALRGCPQSDSDIAPLIDGSGSIIPHDPRMKGFSVSTMQLKKSKTLPALMOYS 193
QY 194 EEPRIHFTKFEFONNPNRSLVKPTQLLGHTRHATGIRKVVRELPNITNGARKNAKIL 253
DB 61 EDPYHFTFNDFKNPSPKLLVRFIRQLLGHTRHATGIRKVVRELPHSKSGARENALKIL 120
QY 254 VITDGEKFGPLGVEDVIPADREGVIRYVIGDAPAFSKSRQELNTIASKPPRDRHV 313
DB 121 VITDGEKFGPLGVEDVIPADREGVIRYVIGDAPAFSKSRQELNTIASKPPRDRHV 313
QY 314 QVNFEALKTQNLREKIPAIETGTGTGSSSSPEHEMSQEGFSAITSGNPLSTVGSYD 373
DB 181 QVNFEALKTQNLREKIPAIETGTGTGSSSSPEHEMSQEGFSAITSGNPLSTVGSYD 373
QY 374 WAGGVFLTSKSTFIMNTRVDSMDNDAYLGAAAIILNRVQSIVLGAPRYCHGLV 433
DB 241 WAGGAFLHMPKQVRFINTTRVDSMDNDAYLGAAVILNRVQSIVLGAPRYCHGLV 433
QY 434 MFRONTGHWESNANVKTQIGAFGASLCSVDVDSNGSDTLVLIGAPHYTQTRGQVSV 493
DB 301 MFKNSGAWKKNADIKGSIQIGSYFGASLCSVDVNRDSSDLVLIGAPHYTQTRGQVSV 493
QY 494 CPLPRGQARWQCDVLYGEGQPGWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAV 553
DB 361 CPLPQG-RAXWQCRVILCGEGHPWSRFGAALTALGDVNGDKLTDVAIGAPGEDNRGAV 419
QY 554 YLFGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMGLVDLTVGAQGHVLL 613
DB 420 YLFGTSELGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMGLVDLTVGAQGHVLL 479
QY 614 RSQVPLRVKATMBFNPREVARNVPECNDQVVGKEAGEVRVCLHVOKSTRDLREGQIQS 673
DB 480 RSQVPLRVKATMBFNPREVARNVPECNDQVVGKEAGEVRVCLHVOKSTRDLREGQIQS 673
QY 674 VTYVDLALDSGRPHSRVAFNETKSTRTQVGLGTQTCETLKLQLPNCIBDPVSVILR 733
DB 540 IITVDLALDSGRPHSRVAFNETKSTRTQVGLGTQTCETLKLQLPNCIBDPVSVILR 733
QY 734 LNFSLVCTPLSAPGNLPUVAEDAQALFALPFPEKNCNDNICODDLSITFSFMSLDC 793
DB 600 LNFSLVCTPLSAPGNLPUVAEDAQALFALPFPEKNCNDNICODDLSITFSFMSLDC 793
QY 794 VVGPREFNTVTVVRNDGDSYRTQVTFPFLDLSYRKVSTLQNRSQRSQRLACESASS 853
DB 794 VVGPREFNTVTVVRNDGDSYRTQVTFPFLDLSYRKVSTLQNRSQRSQRLACESASS 853
```

Db	660	VVGPRDLKVTLTVRNCGESYRTQVTFEFPSPDLSYRKVTSQNRQSRLACESDVS	719
Qy	854	TEVSGALKSTSCSINHPIFFPENSEVTENITPDVDSKASLGNKLLKANVTSENMPRTNK	913
Db	720	TEESTALKSTSCSINHPIFFPNSNSEVTENITPDVDSKASLGNKLLKANVTSENMPSSNK	779
Qy	914	TEFOLELPVKYAVYVMTVTSQVSKYLNFTASEVTSRMQHOYQVSNLQBSLPISLVFL	973
Db	780	TEFOLELPVKYAVYVMTVTSQVSKYLNFTASEVTSRMQHOYQVSNLQBSLPISLVFL	839
Qy	974	VPVRLNQTAVIWDPRQVTFSENLSTCHTKERLPSHSDFLABLRKAPVVMCSIAVCORIQC	1033
Db	840	VPVELNRVTWQDPQVTFSSQNLSSCSSTEEIGPDSDFLEKQKTPVLNCSIAVCQKIQC	899
Qy	1034	DIPFPGIQEENATLKNLSP	1054
Db	900	DIPFPGIQEELKVTLKNLSP	920

RESULT 4

Q81VA6

ID

Q81VA6

PRELIMINARY; PRT; 1169 AA.

AC

Q81VA6;

AC

DT

01-MAR-2003 (TrEMBLrel. 23, Created)

DT

DT

01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT

DT

01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DT

DE

Integrin, alpha X (Antigen CD11C (p150), alpha polypeptide).

OS

OS

Homo sapiens (Human).

OC

OC

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC

OC

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX

OX

NCBI_TaxID=9606;

RN

RN

[1]

RN

RP

SEQUENCE FROM N.A.

RC

RC

TISSUE=Blood;

RA

RA

Strausberg R.;

RL

RL

Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.

DR

DR

EMBL; BC038237; AAH38237.1; -.

DR

DR

GO; GO:0008305; C:integrin complex; IEA.

DR

DR

GO; GO:0004895; P:cell adhesion receptor activity; IEA.

DR

DR

GO; GO:0007160; P:cell-matrix adhesion; IEA.

DR

DR

InterPro; IPR000413; Integrin_alpha.

DR

DR

InterPro; IPR002035; VWF_A.

DR

DR

Pfam; PF01839; FG-GAP; 3.

DR

DR

Pfam; PF00357; integrin_A; 1.

DR

DR

Pfam; PF00092; vwa; 1.

DR

DR

PRINTS; PR01185; INTEGRINA.

DR

DR

PRINTS; PR00453; VWFADOMAIN.

DR

DR

SMART; SM00327; VWA; 1.

DR

DR

PROSITE; PS00242; INTEGRIN ALPHA; 1.

DR

DR

PROSITE; PS00234; VVFA; 1.

DR

DR

SEQUENCE 1169 AA; 128521 NW; A17B484FEFC79EB6 CRC64;

DR

Qy	245	ARNAFKILVITDGEKFGDPLQYEDVIPADREGVIRYVIGVDGADPRSEKSSQELMTIA	304
Db	246	ARDDAKILIVITDGEKGGDSLDYKQVIPNADAAGIIRYAIGVGLAFQFNRSNWKELNDIA	305
Qy	305	SKPDRHVFQVNNFEALKTIIQNOLREKI PAIEGTOTGSSSSSFEHMSQEGFSAAITSNGP	364
Db	306	SKPSQEHIFKVEFDLKDILQNOLKEKI PAIEGTETITSSSSFELEMAQBGFSAVFTDGP	365
Qy	365	LLTSVGSYDWAGGVFLYTSKBKSTFINMTKVDSDMDADVILGYAAAIIILNRVQSGLVLGAP	424
Db	366	VLGAVGSFTMSGAFLYPPNMSPTFINMSQENTMDRSDYLSYSTELAKMKGVSQSLVLGAP	425
Qy	425	RYOHIGLVAMFRONTGWNESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVLGAPHYYE	484
Db	426	RYOHTGXAVITQVSRQWRKAEVGTQIGSYFGASLCSVDVDSGSDSTDLVLGAPHYYE	485
Qy	485	QTRGGQVSVCPFLPRQARQWQCDVILYEGQGPWGFGAALTIVLGVNVDGKLTIDVAIGAP	544
Db	486	QTRGGQVSVCPFLPRGWR - RWCDAVILYEGQHPWGFGAALTIVLGVNVDGKLTIDVVI	544
Qy	545	GEEDNRCAYLPHGTSGSGISPSHSORIASKLSPLQYFGQSLSGGDLTMDGLVDLTV	604
Db	545	GEKENRCAYLPHGTVLGPGSISPSSHSORIASGUSLSLQYFGQALSQGDUTDGLVDLAV	604
Qy	605	GAQCHVLLRSQVLRVYKALMEENPREVARNPFCNDVVYKGEAGEVRVCLHVQKSTRD	664
Db	605	GARGQVLLATRPVLWVGUSMQIFPAEIPRSAFECHREQVVSQETVLQSNCLIVDKESKN	664
Qy	665	RLBGGIQSVVTDLALDSGRPHSAVFNETKNSRTOQTVLGLTQCTETLKLQLENCIE	724
Db	665	LLGSRDLQSSVTLTDLALDPGRLSFRATFOETKRSISRVRVLGKAKHCFNFIILLSCVE	724
Qy	725	DPVSPVILRNPLNPLVGTPLSAFGNLAPVLAEDAQRLPTALPPPEKNCNGDNIQDDLSIT	784
Db	725	DSVTPITILNPLTVGKPLLAFNRLPMLAADAQRYFTASLPEKNCGADHIQODNLGIS	784
Qy	785	FSPMSLDCLVWGPRBENVTVTVRNDGDSYRQTVTPFPPLDLSYRKVSTLQNKORSQSW	844
Db	785	FSPPGLSKSLVAGSNLELNAEVMVWNDGDSYGTVTTFSHPAGLSYRYVARGQKQQLRSL	844
Qy	845	RLACBSASSTEVSGALKSTSCSINHPIFPENSESVTNTITFDVDSKASLGNKLLKANVTS	904
Db	845	HLTCDGAPVG - SQGTWTSISCRINHILPRGGAQITFLATFDVSPKAVLGDRIILLTANVSS	902
Qy	905	ENNMPTNKTEPQLELPVKYAVTMVYVTSHGVSSTKYLNFTAS - ENTGRVMOHQYQVSNLQG	963
Db	903	ENNTPTSKITFQLELPVKYAVTVVYSSHEQTKYLNPSGESBEKSHVAMERYQVNNLQG	962
Qy	964	RSIPISLVPLVRLNYWINDRPOVTFSENLSSTCHTKERLPSHSDPLAELKAPVUNC	1023
Db	963	RDLPVSIINFVWVELNQBAVMVMDVSHFPQNPSLRCSSEKIAPPASDFLAHQNPVLDG	1022
Qy	1024	SIACVORIQCIDIPIFGIQBEFNATLKGNISFDWVYIKTSHNHLIVSTAETILFNDVSFTLL	1083
Db	1023	SIAGCLRFRCDVPSFSVQELDPTLKGNLSPGVMVQILQKKVSVSVVAEITFTDSVYSQL	1082
Qy	1084	PGGCAVRSOTETKVPFPFVNPDLPLIVGSSVCGILLLLALITAAALKLGFKKFYQKDMMS	1143
Db	1083	PGQEAFWRAQTITVLEBKIKYHNPTPLIVGSSIGGLLLALITAVLYKVGFPKRYKEMME	1142
Qy	1144	E 1144	
Db	1143	E 1143	
RESULT 5			
Q9QXH4			
ID	Q9QXH4	PRELIMINARY;	PRT; 1169 AA.
AC	Q9QXH4;		
DT	01-MAY-2000	(T=EMELrel. 13, Created)	
DT	01-MAY-2000	(T=EMELrel. 13, Last sequence update)	
DT	01-JUN-2003	(T=EMELrel. 24, Last annotation update)	
DE	Leukocyte adhesion glycoprotein p150,95 alpha integrin subunit.		

GN ITGAX.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
QX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Huang X., Goreki K., Tong C., Rattis F.-M., Tseng S.-Y., Pardoll D.,
RA Tsuchiya H.;
RT "Isolation of Genes Selectively Expressed by Dendritic Cells.";
RT Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF211864; AAF23492.1; -;
DR HSSP; P11215; 1BHQ.
DR MGD; MGI:96609; Itgax.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:007160; P:cell-matrix adhesion; IEA.
DR GO; GO:0007229; P:integrin-mediated signaling pathway; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PRQ1185; INTEGRINA.
DR PRINTS; PRQ0453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; vwa; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS02334; VWFA; 1.
KW Integrin.
SQ SEQUENCE 1169 AA; 129150 MW; C616412033C219A6 CRC64;

Query Match 56.3%; Score 3352.5; DB 11; Length 1169;
Best Local Similarity 57.1%; Pred. No. 116-234; Mismatches 168; Indels 19; Gaps 7;
Matches 657; Conservative 168; Mismatch 234; Indels 19; Gaps 7;

QY 6 LLLTALTALCHGNLDNTENAMTQENARGFGQSVVQLQGSRRVVGAPQRTVAANRQSGSLYQ 65
DB LLLLGFSVCLGNLDAEKLTHFMDGAERGHVLYQDSSVVVVGAPKEIKATNOIGLYK 68

QY 66 CYSTGSCPIRLQVPVEAVNMSLGLSAAATSPOLLACGPTVTCSTENYVKGCLPL 125
DB CGYHTGNCPEPISLQVPPPEAVNTSLGLSAAATSPNLLACGPTVHTCRENTLYLGLCL 128

QY 126 FGSNLRLQPKPEALRGCPQEDSDIAFLIDSGSIIPIHDFRMKEFVSTVMEQLKSKT 185
DB LSSSFQKS-QNPTAQCEPKQDQDIVFLIDSGSISSTDFEMLDVFAVMSQLQRST 187

QY 186 LFSLMQYSSEPRHFTFKFQNNPNRSLVKPITQLLGRHTATGIRKVVRELFMTNGA 245
DB RFSLMQFSDYFRVHTFTRNFPISTSSPLSLGSRQLRGVYTTASAKHVITELFTQSGA 247

QY 246 RXNAPKILVITDGEKFGDPLGYEDVPEADREGVIRYVIGYDAPRSEKSPQELNTIAS 305
DB RQDATKVLIVITDGRKQGNLSYDSVIPMAEASIIRYAIGVGKAPYNEHSHQELKAIAS 307

QY 306 KPPRDEHVFQNNFEALKTIQNLREKIPAEIGTQTGSSSFHEMSQEGFSAATSNGLP 365
DB MFSHEVSVFENFDALKDIENLKEKIPAEIGTETPSSSTFELEMSQEGFSAVTPDGPV 367

QY 366 LSTVGSYDWAGGVFLYTSKEKSTFNTNTRVDSMDADYLGAAAILNRVQSVILGAPR 425
DB LGAVGSFSSGGAFLYPSNRPFTINMSQENMDRDYLGSTALFAPKGVHSLILGAPR 427

QY 426 YQHIGLVAFRQNTGHWESNANVKGITQIGAYFGASLCSDVDNSGSDTLVILGAPHYEQ 485
DB HQHTGKVITPQBSRHRPKSEVRGTQIGSYFGASLCSDVMDRDGSDTLVILGVPHYEH 487

QY 486 TRGGQVSCPLPRGQRARQCCDAVLVBGQGPWGRFGAALTVLGVDNGDGLTDVAIGAPG 545
DB TRGGQVSCVCPMP-GVGSRRHCGTTLFGEQGPWGRFGAALTVLGVDNGDGLADVAIGAPG 546

QY 546 BEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPLRQYFGQSLSGGQDLTMDGLVDLTVG 605

DB 547 EENRGAIVYTHGASRQDIAPSPQRISASQIPSRIOYFGQSLSGGQDLTRDGLVDLAVG 606
QY 606 AQGHVLLRSQPLRVKAIMENPREVARNYFECNDQVVKGEAGEVEVCLHVQKSTDR 665
DB 607 SKGRVILLRTPILRVSTVHTFPAISRSVFQEQVAPQOTLSDATVCLHIESPKIQ 666
QY 666 LREGIOISVWTYDLALDSGRPHSAVFNETKSNTRRQTVLGLTQTCFLLKQLPNCIED 725
DB 667 L-GDLRSTVTFDLALDHGRSLSTRAIPKTKTRALTRVKTLLGNKHCESVKLLPACVED 724
QY 726 PVSPIVLELNFSLVGTPLSAFNGNLPALEAQLFTALFFERKNCNDNTCODLSITF 785
DB 725 SVTPITLELNFSLVGVPISSILQNLQPLAVDDQYFTASLFFERKNCNCGADHICODDLSVVF 784
QY 786 SFMSLDCLVWGSPREFNVTVTRNDGEDSYRTQVTFPPPLDLSVRKVVSTLQ----- 836
DB 785 GPPDLKTLVGSDELNVDTVSNDEGDSYGTVTTLFYPVGLSFRVAEGQVFLRKKEDQ 844
QY 837 --NORSQSRWLACESASSTSVGALKSTCSINHPIPPENSEVTFNITFDVDSKASLON 894
DB 845 QWRRGQSHLMCD--STPDRSQGLMSTSCSRHVPFRGSGQMTFLTVFVSPKAEIGD 902
QY 895 KLLKANVTSENNPRTMKTBFQLELPVKYAVYVWVTSHGVSTKYLNFTASE-NTSRVMO 953
DB 903 RELLARVGSNNVGPCTPKTFQLELPVKYAVYVWVTSHGVSTKYLNFTASE-NTSRVMO 962
QY 954 HQYQVNLGQSLPISLVLVPLRVINQTVIWRFOVTFSENLSTCHYKRLPSHSDPLA 1013
DB 963 HRFQVNLGQSLPISLVLVPLRVINQTVIWRFOVTFSENLSTCHYKRLPSHSDPLA 1021
QY 1014 ELRQAVVNCISIAVQRIQCDIPFGIOEENATLKGNLSPDWYIKTSHNLLIVSTARI 1073
DB 1022 HQKSPVLVDCSLADLHURCDIPSLGILDELDELFLKGNLSFGNITSQTLQKVVLLSEAI 1081
QY 1074 LFNDVFTLLPQGAFAVRSQTTETKVEPFEVFNPLVGVSSVGGLLALLALITAALYKLG 1133
DB 1082 TFSVSVSQLPQGAFLRAQTKTVLEMYKVNPNVPLVGVSSVGGLLALLAITAILYKAGF 1141
QY 1134 FKROYKDMSE 1144
DB 1142 FKROYKEMLEE 1152

RESULT 6
Q9QYE7 PRELIMINARY; PRT; 1161 AA.
AC Q9QIE7;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Alpha D integrin.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA O'Brien M.M., VanderVieren M., Kilgannon P.D., Dietsch G.,
RA Gallatin W.M.;
RT "Cloning of rat alpha D, a novel beta 2 integrin.";
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF021334; AAF21241.1; -;
DR HSSP; P11215; 1BHQ.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR GO; GO:0007229; P:integrin-mediated signaling pathway; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.

```
DR Pfam: PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR SMART; SM00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS50234; VWA; 1.
KW Integrin.
SQ SEQUENCE 1161 AA; 126600 MW; 22584913984705E CRC64;

Query Match
Best Local Similarity 55.1%; Score 3282.5; DB 11; Length 1161;
Matches 663; Conservative 158; Mismatches 306; Indels 15; Gaps 10;

QY 5 VLLIT--ALTCHGENLDTENAMTFQENARGGQSVVQLQGSRVVVGAPQEIIVAAANORG 62
DB 6 VILLCGVLASCHGSLNEDVEEIVFREDAASTQIVVQGGSLRVVGAPLEAVANQTR 65
QY 63 LYQDYSTGSCBPIRLQVPEAVNMSLGLSLAATSPQQLACGPTVHQTCSYNTYVKG 122
DB 66 LYDCAPATGWCQPIVLSRPLEAVNMSLGLSLVATNNAQLACGPTAQRACVKNYAKG 125
QY 123 CFLFGSLRQKQKPFALRGCPQEDSDIAFLIDGSGIIPHDFRMKEFFVTVMEQLKK 182
DB 126 CULLGSSL-QTQAVPASMEPCPEQEMDIAFLIDGSGINQDFQAKMDFFVXALMGEPAS 184
QY 183 SKTLPFLMOYSEFRTHFTKFPQNNPNRSLVKPTTQALLGRTHATGIRKVVRLFNIT 242
DB 185 TSTLPFLMOYSNLTKHTFTFTFQNLDPQSLVDPIVQLQGLTYTATGIRTVMEBLFHSK 244
QY 243 NGARKNAFKILVITDGEKFGDPLGYEDVIVPEADREGVIRYVIGCDAPRSKSRQELNT 302
DB 245 NGRSKSAKILLVITDQKRPDLEYSVDVIPAADKAGIIRYAGVDAPOBPTALKELNT 304
QY 303 IASKPPRDHVFOVNPPEALKITQNLREKI FATEGTQTGSSSFHEMSQEGFSAAITSN 362
DB 305 IGSAPPQDHVFKVNGFAALRSIQRLQLEKI PALEGTQSRSSSFQHEMSQEGFSALTSD 364
QY 363 GPLLSTVGSVDAGGVFLYTSKEKSTFIMNTRVDSMDNDAYLGYAAAILNRNVSQVLG 422
DB 365 GPVLGAGVSPSGGAFYFPNTRPFFINNSQENVMDRDSYLGYSTAVAFWGVHSLILG 424
QY 423 APRYQHIGLVAMPFRONTGMWSESNANVKGITQIGAYFGASLCSVDVDSNGSDTLVLICAPHY 482
DB 425 APRHQHTGKVIPTQBARHWRPKSEVRGTQIGSYFGASLCSVDVDSGTDVLICAPHY 484
QY 483 YEQTRGGQSVVCPPLRGQBARWOCDAVLYGEOQPMGRFGAALTVLGDNVNGKLTDTVAIG 542
DB 485 YEQTRGGQSVVFPVP-GVGRWQCEATLHGEQHPMGRFGVALTVLGDVNGDNLDVAIG 543
QY 543 ARCEENRGAVYLPHTGSGGISPSQSORTAGSKLSPRLQYFGOSLGGQDLTMDGLVDL 602
DB 544 APGEERSRGAVYIFHGASRLNMFSPSQRSQVLSRLQYFGOSLGGQDLTMDGLVDL 603
QY 603 TVGAQGHVLLLRQPVLRVKAIMFENPREVARNVFECDQVVGKGEAGEVVRVCLHVQKST 662
DB 604 AVGAQGHVLLLRSLPLKVELSRFAPMEVAKAVYQCWERTPTVLBAGEATVCLTVHKGS 663
QY 663 RDLRGQIQSVVYDIALDSDGPHSRVAVNETKNSRTQTVLGTQTCETILKQLPNC 722
DB 664 PDLL--GNVQGSRYDLALDPGLISRAIFDETCKNTLTGRKTLGHDCHETVKLLPDC 721
QY 723 IEDVPSPVILRLNFSLVGTPLSAFAGMLRPVLAEDAQLFTALFPFPEKNCNMDNICDDLS 782
DB 722 VEDAVSPITILRLNFSLVDSASP-RNLHPVLAVGSDPHITASLPFEKCKQELLCEGLG 780
QY 783 ITFSFMSLCLVVGGRPREFNVTVTVNDGEDSVRTQVTFPPFLDLISYRKVSTLONORSR 842
DB 781 ISFNFSGLQVLVVGSGPELTVTVTVNVEGSDYGLVKEFYFAGLSYRRTVGTQ-QPHQY 839
QY 843 SWLACESASTSVGALKSTCSINHPIFENSEVTENITPDVDSKASLGNKLLKANV 902
DB 840 FLRLACEAPEAQED--LRSSCSINHPIFREGAKTFTMTITPDVSYKATFLGDELLLRAXA 897
```

```
QY 903 TSENNMPTNKTROLELPYKYAVYVWVTSHGVSSTKVLNFTASENTSR-VNQHGYOVSNL 961
DB 998 SSENKPGDTNKTAFQLELPVAYITVTLISROEDSTNHNVSSTSGGRRQEAHRYRVNL 957
QY 962 QORSLPISLVELFVRLNQTVIMDRPQVTFSENLS--TCTKERLPSSHDFLABLRKAPV 1020
DB 958 SPLKLAUVNFWPVLNGVAVMD---VTLSPPAQGVSCVSMKPPQNPDLTQIRRSV 1014
QY 1021 VNCSTAVCORIQCDIPFFGIOEENATLKGNLSPDWIKTSHNLLIIVSTAELFNDSVF 1080
DB 1015 LDCSIADCLHFRCDIPSLDIQDELDFILRGNLSRGWSQTLQEKVLLVSEAEITFDSVY 1074
QY 1081 TLLPGQAFVRSQSTQETKVEPEVENPLPLIVGSSVGGILLIALITAAALYKLGFFKRYKD 1140
DB 1075 SOLPGQAEFLRAQVSTTELEVYVYEPFLVAGSSVGGILLIALITVVLXKLGFFKRYKE 1134
QY 1141 MW 1142
DB 1135 ML 1136

RESULT 7
Q9WTV4 PRELIMINARY; PRT; 1161 AA.
AC Q9WTV4;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Integrin alpha L.
GN ITGAL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI TaxID=10090;
RN [1]
RC SEQUENCE FROM N.A.
RA STRAIN=DEA/23; TISSUE=Spleen;
RL Submitted (May-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF065902; AAD25885.1; -.
DR HSSP; P20701; 1LFA.
DR MGD; MGI:96606; Itgal.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; P:cell adhesion receptor activity; IEA.
DR GO; GO:0007180; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VMP_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS50234; VWA; 1.
SQ SEQUENCE 1161 AA; 128240 MW; 86B102F7B2098431 CRC64;

Query Match
Best Local Similarity 25.9%; Score 1545; DB 11; Length 1161;
Matches 407; Conservative 219; Mismatches 462; Indels 108; Gaps 39;

QY 1 MALRY---LLLTALT---CHGFNLDTENAMTFQENA-RGFGQSVVQLQGSRVVVGAPQ 52
DB 1 MSFRIAGPRLLLLGLQLPFAKAWSYNLDTRPTQSLAQAGRHFGYQVLQIEDG-VVVGAPG 59
QY 53 EIVANQRGSLYQCDYSGSCPEIRLQVPEAVNMSLGLSLAATSPQQLACGPTVHOT 112
DB 60 E---GDNTGGYHCTSSSEFCQPSLH--GSNHTSKYLGMTLATDAKGSLLACQPLGRT 115
QY 113 CSENTYVVKGLCFGLFGSNLRQKQFPPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEF 172
```

Db 116 CDQNTYLSGLCYLPQSLEGPMLQNRPAVQCMKGVLDVLFVFDGSGSLDRKDFKILF 175
 Qy 173 VSTVMEQLKSKTFLSLMOYSEFRHFTPEF-QNNPNRSLVKETITOLLGRTHATGI 231
 Db 176 MKDVMEKLSNTSQFAVQSTDCRTEFTLDVYKQNKPNPDVLLGVSQPMFLITNTFRAI 235
 Qy 232 RKVVRELFNITNGARKNAFKILVITDGRKFGDPLGYEDVPEADREG-----VIRYV 284
 Db 236 NYVAVHFKESGARPDAKVLVIITDG-----EASDKGNISAADITRYI 281
 Qy 285 IGVGDAPFRSEKROELNTIASPPDHVFOVANNPEALKTIQNLREKIFAIGTQTGSS 344
 Db 282 IGIKHFVSQKQTHIFASEPVEFVKILDTFELKDLFTDLQRIYVAIGTNRQDIT 341
 Qy 345 SPEHEMSQEGFAAITSNGPLLSVGSYDWAGGVF-LYTSKEKSTFINMTRVDSMDNDAY 403
 Db 342 SPNMELSSGISADLSKGHAVGAKWAGGFLDLREDLQATFVGOEPLTSDVRGY 401
 Qy 404 LGYAAA-IILNRVQSLVGLGAPRYOHIGLVAMFR--QNTGMWESNANVKGTQIGAVFGAS 460
 Db 402 LGYTVAMTSSRRLPLAAGAPRYOHVGVLLFQAPDAGGRWNQTKIETGTQISYFGE 461
 Qy 461 LCSVDVDSNGSTDLVLIGAPHYETOTRGQSVCPPLRGQARWQCDVLYGSGQGPWCR 520
 Db 462 LCSVDVDSNGSTDLVLIGAPHYETOTRGQSVCPPLRGQARWQCDVLYGSGQGPWCR 520
 Qy 521 FGAALTVDGVNGDKLTDVAIGAPGEDNRGAVYLFHGTSGSGISPSHQIAGSKLSPR 580
 Db 519 FGAALTVDGVNGDKLTDVAIGAPGEDNRGAVYLFHGTSGSGISPSHQIAGSKLSPR 580
 Qy 581 LQVFGSLSGQDLTWGLVLDVTVGQCHVLLRSQPLVLRKAI MEFNPREVARNVFECN 640
 Db 576 IRWFGSRHGVKDLGGDLADVVGVGPRVVLSSRPVVDVVTLSFSEEPVHVECS 635
 Qy 641 DVVKGKEAG-EYRVGLHGVKSTEDRLREBQIOSVVTYDIALDSGRPHRAVFNETKST 699
 Db 636 YSAREEQKHGVKLCACRIKPLTPQ--FQGRLLANLSYTLQLOGRHMRSGFLPDDGSHEL 693
 Qy 700 RQOTVGLGLTQTCETKLQPLNCEBDSVIVRLNFSLV---GTPLSAFGR-LRPVLAE 755
 Db 694 SGNSTITP-DKSLDFHFPFICIQDLISPIVNSLNFSLLEBEGTPRQXGRAMQPIRP 752
 Qy 756 DAQRLFTALPFPFKNCNDNIQDLSITPFSMSLCLVVGGE-----REFNVTVVRN 809
 Db 753 SIHTV-TKEPPEKNGCEDKCEANLTLSPARS-----GFLRLWSSASLAVETLSN 804
 Qy 810 DGEDSVRTQVTFPFLDLSTVRKYSTIQNSQSRWRLACBSASTEVSGAL-KSTSCSIN 868
 Db 805 SGEDAYVWRLDLPFRGLSPKRVEMLO---PHGRMPVSCBEL--TEGSSLLTKCNYS 859
 Qy 869 HPPIFENSEVTNITFDVDSKASLGNKLLKANVTSEN-NMPTNKTETQLELPVKVAVY 927
 Db 860 SPIFKAGQVSLQVNFLLNSWEDFVGLNGTVHCENENSSLOEDNSAATHIPVLYPN 919
 Qy 928 MVYTHGVSTKYLNFTASENTSVMQHQYQV---SNLQSRSLPISLVFLVPLVNLQVITW 984
 Db 920 ILTKQENSTLXISFTPKGPKTQOVQHVQVRIQPGAYDHNM-TLEALVGP-----W 972
 Qy 985 DRPQ--VTFSENLS---TCHTKE-RLPSHSDFLAELRKAPVNCIAVQRIQDIPF 1037
 Db 973 PHSEDPITYWSDTDLVTCRSEDLLKRPSE---AEQPCPLPV-----QFRCPITV 1021
 Qy 1038 FGIQEBFNATLKNLSPDWIKTSHNLLIVSTAEILFNDVSFTLLPFGQAFVRSQETK 1097
 Db 1022 ---RRETLIQTVGVLSKEIKAS-STLSLCSLSVSFNSSKHFLVGSKA-SEAQVLVK 1076
 Qy 1098 VEFPEVNPPLVIGSVGGLLILALITAYLYKLGFPKRYKQDM-SEGGPQAE 1152
 Db 1077 VDLIHEKEMLVVYLSGIGGLVLLFLFLALYKVGFPKRLKEXMEADGVPNGSP 1132

RESULT 8
 Q9R200

ID Q9R200 PRELIMINARY; PRT; 1160 AA.
 AC Q9R200;
 DT 01-MAY-2000 (TREMELrel. 13, Created)
 DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
 DE 01-JUN-2003 (TREMELrel. 24, Last annotation update)
 DE Integrin alpha L.
 GN ITGAL.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN 113
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Spleen;
 RA Ma R.Z., Teuscher C.;
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF065901; AAD25884.1; --
 DR HSSP; P20701; 1LFA.
 DR MGD; MGI:96606; Itgal.
 DR GO; GO:0008305; C:integrin complex; IEA.
 DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
 DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
 DR InterPro; IPR000413; Integrin_alpha.
 DR InterPro; IPR002035; VWF_A.
 DR Pfam; PF01839; FG-GAP; 3.
 DR Pfam; PF00357; integrin_A; 1.
 DR Pfam; PF00092; vwa; 1.
 DR PRINTS; PR01185; INTEGRINA.
 DR PRINTS; PR00453; VWFADOMAIN.
 DR SMART; SM00191; Int_alpha; 5.
 DR SMART; SM00327; VWA; 1.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 DR PROSITE; PS02324; VWFA; 1.
 DR PROSITE; PS02324; VWFA; 1.
 SQ
 Query Match 25.8%; Score 1534.5; DB 11; Length 1160;
 Best Local Similarity 33.9%; Pred. No. 2,7e-102;
 Matches 405; Conservative 218; Mismatches 465; Indels 107; Gaps 38;
 Qy 1 MALRV-----LLTALT---CHGFNLDTENAMTFENA-RGFGQSVVQLGSGRVVVGAPQ 52
 Db 1 MSFRIAGPRLLLGLQLFAKWSYNLDRPTQSLAQAGRHFGYQVLIQEDG-VVVGAPG 59
 Qy 53 EIVAAQNGSLYQCYSTGSCBPIRLQVEVAVNMSLGLSLAATSPQOLLACGPTVHOT 112
 Db 60 E---GNTGGLYCHTSSFEFCQPVSLH-GSNHTSKYLGWTLATDAKGSLLACDPLSRT 115
 Qy 113 CSNTYVKGCLPFGSNLRQOPKPEALRGCPQSDSDIAFLIDSGSIIIPHDPRMKPEF 172
 Db 116 CDQNTYLSGLCYLPQSLEGPMLQNRPAVQCMKGVLDVLFVFDGSGSLDRKDFKILF 175
 Qy 173 VSTVMEQLKSKTFLSLMOYSEFRHFTPEF-QNNPNRSLVKETITOLLGRTHATGI 231
 Db 176 MKDVMEKLSNTSQFAVQSTDCRTEFTLDVYKQNKPNPDVLLGVSQPMFLITNTFRAI 235
 Qy 232 RKVVRELFNITNGARKNAFKILVITDGRKFGDPLGYEDVPEADREG-----VIRYV 284
 Db 236 NYVAVHFKESGARPDAKVLVIITDG-----EASDKGNISAADITRYI 281
 Qy 285 IGVGDAPFRSEKROELNTIASPPDHVFOVANNPEALKTIQNLREKIFAIGTQTGSS 344
 Db 282 IGIKHFVSQKQTHIFASEPVEFVKILDTFELKDLFTDLQRIYVAIGTNRQDIT 341
 Qy 345 SPEHEMSQEGFAAITSNGPLLSVGSYDWAGGVF-LYTSKEKSTFINMTRVDSMDNDAY 403
 Db 342 SPNMELSSGISADLSKGHAVGAKWAGGFLDLREDLQATFVGOEPLTSDVRGY 401
 Qy 404 LGYAAA-IILNRVQSLVGLGAPRYOHIGLVAMFR--QNTGMWESNANVKGTQIGAVFGAS 460
 Db 402 LGYTVAMTSSRRLPLAAGAPRYOHVGVLLFQAPDAGGRWNQTKIETGTQISYFGE 461
 Qy 461 LCSVDVDSNGSTDLVLIGAPHYETOTRGQSVCPPLRGQARWQCDVLYGSGQGPWCR 520

Db 462 LCSVDLDQDQGEALLICAPLFFCQGRGVFTY---QRQSLFEMVSELOQDGPVPLGR 518
 QY 521 FGAALTALVGLVNGDKLTDVAIGARBEEDNRCANVILFHTGSGISPSHSGRIAGSKLSR 580
 Db 519 FGAALTALVGLVNGDKLTDVAIGARBEEDNRCANVILFHTGSGISPSHSGRIAGSKLSR 575
 QY 581 LOYFQSLGGQDUTMDGLVLTGACQHVLLRSQPLVAKVIMENPREVARNVFECH 640
 Db 576 IRWEGRIHGVKGLGDBELANVVGPRVVLSSRPVVDVTELSPSPREIPVHEVCS 635
 QY 641 DOVVKKEAG-EVRVCLHVOKSTRDLRREGIOISVYDLDLSDGRPHSRVAVNEKNST 699
 Db 636 YSAREEOKHGVKACPRKPLTPQ--FOGRLANLSYTLQLDCHRMRSRGLPFDGSHL 693
 QY 700 RROTVLGLTCTCTBLKQLPNCIEDPVSPVLRNLSLV---CTPLSARFN-LRPVLAE 755
 Db 694 SGNTSITP-DKSCLDLFFHPICQIDLISVNSLNFSLLEBETPRDQGRAMQILRP 752
 QY 756 DAQSLFTALPFFKNCNNDIQQDLSTTFMSLDCLVVGGP-----REFNVTVVRN 809
 Db 753 SIHTV-TKEIPFKNCGEDKCEANLTLSPPARS-----GRLMSSASLAVETLSN 804
 QY 810 DGDSYRTOVTFPPPLDLSYRKSTLQNRQSRWSRLACESASTEVSGAL-KSTSCSN 868
 Db 805 SGDAYVWRDLDPRLGSLFRKVMELQ---PHSRMPVSCBEL--TEGSSLLTTLKCNVS 859
 QY 869 HPIPPENSEVTNITFDVDSKASLGNKLLKANVTSEN-NMPRTNKTFFOLELPVYAVY 927
 Db 860 SPIFKAGQVSLQVMTLLNSWEDFVNLGTVHCENENSLQEDNSAATHIPVLPVN 919
 QY 928 MVTSHGVSTKYNLTASENTRVMOHQVSNLQORSLPSLVFLVPLNQTWDRP 987
 Db 920 ILTKQENSTLYSTFKGPKTQVQVQV---RIQSAVDHNPET-LEALVGVPRP 973
 QY 988 Q----VTFSENLS--TCHTKE-RLPSSHSDFLAELRKAPVNCSTIAVCQRIQCDIPFF 1038
 Db 974 HSEDLIYTVSVQDPLVCHSBDLKPS-----SEABPCLGV-QFRCPVFP- 1020
 QY 1039 GIOEFNATLKNLSFWYIKTSHNLLITVSTABILFNDVFTLLPQOGAFVRSQETKV 1098
 Db 1021 --RWEILIQVGTVELSKEIKAS-STLSLCSLSVSNFSSKPHLYGSKA-SEAQVLKV 1076
 QY 1099 EPFEVPNPLPLIVGSSVGGILLALITAAALYKLGFFKQVQKMW-SEGPPGNAEP 1152
 Db 1077 DLIHEKEMHVVVLSGIGGLVLLFLIPLALYKVGFFKELKEXHEADGGVPNGSP 1131
 RESULT 9
 Q98TF1
 ID Q98TF1 PRELIMINARY; PRT; 1196 AA.
 AC Q98TF1
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Coli-1.
 GN CIA1.
 OS Cyprinus carpio (Common carp).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Cyprinus.
 OX NCBI_TaxID=7962;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Peritoneal exudate cells;
 RA Kimura M., Nakao M., Miura C., Fujiki K., Yano T.;
 RT "Molecular cloning of a leukocyte integrin from the common carp.";
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 DR ENBL; AB048536; EMBL39134.1; -.
 DR HSSP; P20701; ILFA.
 DR GO; GO:0008305; C:integrin complex; IEA.
 DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
 DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
 DR GO; GO:0007160; P:cell-matrix adhesion; IEA.

GO; GO:0006508; P:proteolysis and peptidolysis; IEA.

InterPro; IPR001969; Aspartic-type endopeptidase AS.

InterPro; IPR000413; Integrin_alpha.

InterPro; IPR002035; VWF A.

Pfam; PF01839; PG-GAP; 3.

Pfam; PF00357; Integrin_A; 1.

Pfam; PF00092; vwa; 1.

PRINTS; PRO1185; INTEGRINA.

PRINTS; PRO0453; VWFADOMAIN.

SMART; SM00191; Int_alpha; 5.

SMART; SM00327; VWA; 1.

PROSITE; PS00141; ASP_PROTEASE; 1.

PROSITE; PS0234; VWF_A; 1.

SEQUENCE 1196 AA; 132477 MW; 9369C807E7DCA53B CRC64;

Query Match 23.8%; Score 1419; DB 13; Length 1196;

Best Local Similarity 31.3%; Pred. No. 7.3e-94;

Matches 377; Conservative 228; Mismatches 455; Indels 144; Gaps 39;

3 LRVLLLTALTLCGFNLDTEENAMTFQENARG-FGQSVVQLQ-GSR--VVVGAPQBIIVAN 58

18 LGILLFMWASLGEAFNIDTEHPLRNGAPEDFFGVYVQTEFGNRKQIIVGAPLE---GN 74

59 QRGSLYQDYSTGSCPEIRKQVP-----VAVNMSLGLSLAATSPOLLACQPTVHQTGS 114

75 STGEMYSCTADLQSCQ--RLQRPSESVRFGKSAVSAALTS-----CSPIYPHECD 126

115 ENTYYKGLCFLFGSNLRQOPKQPFALRCQPEDSDIAPLDGSGSIIPHDERRMKFVS 174

127 GNSYLVGVYQSSSL-QAVSNFTAYQECSEKENVLFLPDGSSSMKTVDEMKNFK 185

175 TYNEQKQKKTFLSLMQISREPRHIPTKBPONNPNRELVKPTQLLGRTHATGIRKV 234

186 DIMKLSNSIIFAFAVQFSTDVRTVDFNDYQSGSAEKLME-THMSLTNTHKAIDVI 244

235 VRELEN-INGARKNAFKLVITDGEKFGDPLGYED--VIPEADREGVIRVYVIGVDAP 291

245 LKNLNSMLSGADSAQALVITD---GDPSDNDVNVLLKCKDEQNLRIIIVG--- 297

292 RSEKSRQELNTIASKPPRDHVFQVNFALKTIONQLRKIFAIETGTQTSSSSPEHMS 351

298 --KVDLIELTQLASEPKRNTFVKDYGLGLLDNLQKKIYNIEGSDVAQDRDKELS 355

352 QSGFSNAITSNGLPLSTGVSVDWAGVFLYTSKEKSTFINMTFRVSDNM-DAYLVGAAI 410

356 QSGFSVTVQESVIVGSGVNDWRGALYEVTC-SGSDFAETIIDPAVNGSIMGYSTV 414

411 ILERNVQSLVGLGAPRYQHIGLVAMFRONTGMWESNANVKGTOIGAYFGASLCSVDVDSNG 470

415 GMRGVSLLLFSGAPRAEHTGLVTLTKQNTVTVTSNNGEQIGSYFGASLSDVDSDG 474

471 STDVLVIGAPHYEQTGQGVSVCLPRQORARWOCDAVLYCEQ-----GQPKGFGAAL 525

475 DSDFLVVGAPLFYQSQ-----PRTEGRVYVYSLSEBQYFQKTLNVVSQSTTGFAASV 526

526 TVLVGVNGDKLTDVAIGAPGEDNREGAVVLFHTGSGISPSHS-QRIAGSKLSRLOVF 584

527 ASLKLNGDGLSDVAVGAPLE--NEGVVYVILGDRTHGINPELTQRISQSVLPGLQOF 584

585 QGSLGGQDLTMDGLVLTGVAQGHVLLRSQPLVAKVIMENPREVARNVFECHQDV 644

585 GVSLTGQMDMNDNLTDVIGAGQGVILKARFVMSVAQLSPSPKBSISLNTFPCGS-- 642

645 KGKEAGEVRVCLHVOKSTRDLRREGIOISV--VTYDLALDSCRPHSRVAVNETKNSTR-- 700

643 NAFNANFLTSCFTVETSS---TGSLEKLVNSLVNVDVVRGMSRGFFDQSSVSSRTL 699

701 RQTVLGLTQTCETLKLQPLNCIEBVPSPVILRLNPSLVGTPLSPAGNLRPLVLAEDAKRL 760

700 QGSLVLDGSSCFNFSIFMLRCVADTVSPKIRNFS--QTEMLS-GNSVAVLDVHSRTE 756

761 FTALPFFKNCNNDNICQDDLSITFSFMSLDCLVVGGPREFNVTVVRNDGSDSRYTQV 820


```

Db 961 TTRVTLTPHNFMSQEVHHSFISSYHQ--LIIMCLNKLHFTSPBLSAVQVTRTGRSLWVC- 1017
Qy 1030 RIQCDIFFPGIOEFNATLKGNL-----SFDWYIKTSHNHLIIYSTHILF 1075
Db 1018 -----VSSISTGELFRSSV--NLMAEAVLQNKVEYSKISFY-EFRDRHVNIS-AELNF 1068
Qy 1076 NDSVFTLLPG-OQAFVRSQSTETKVEPEVFNPLPIVGVSGVGLLLALITAAALYKLGFF 1134
Db 1069 NTSRYNSTGLKNPHSQTVEKVFEVIPPSSRLIVCTGAVGGFFFIILIIILLKCGFF 1128
Qy 1135 KROYKD 1140
Db 1129 KNRNPD 1134

RESULT 11
Q96HBI PRELIMINARY; PRT; 1086 AA.
AC Q96HBI;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DB Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RX [1]
R2 SEQUENCE FROM N.A.
RC TISSUE=Lymph;
RA Strausberg R.;
RL Submitted (May-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; EC008777; AA008777.1; -.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; P:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF018139; PG-GAP; 3.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWF_A; 1.
KW Hypothetical protein.
SQ SEQUENCE 1086 AA; 119223 MW; F6FF2546B8C632P9 CRC64;

Query Match 22.9%; Score 1361; DB 4; Length 1086;
Best Local Similarity 32.0%; Pred. No. 1e-89;
Matches 381; Conservative 188; Mismatches 436; Indels 184; Gaps 38;

Qy 6 LLTALTLCGPF-----NLDENAMTFQ--ENARGFGQSVVQLQSRVVVGAPQEI 54
Db 6 ITVMALLSGFFFPASSYNLDVRGARSFSPFRAGRHFGRVQLQV-GNGVIVGAPGE- 63
Qy 55 VAANORSLYCCDYSTGSCPIRLQVFBVAVNMSLGLSLAATTSPPQLACGTVHTGCS 114
Db 64 --GNSTGSLYCCQSGTCHCLPVTLR-CSNYSYKYLEMTLA--TDP----- 103
Qy 115 ENTVYKGLCLFGLNLRQOPKQPPALRGCPQEDSDIAFLIDGSGSIIPHDFRMKFVS 174
Db 104 -----TDGS----- 107
Qy 175 TVMEQLKSKTFLSLMOYSBEERIHPTFEFQNNPNRSIAVKPITOLLGRTHATGIRKV 234
Db 108 -----ILPAAVQFSYKTFEDFSDYVRKOPDALRKRVKMLLNTFGAINTV 157
Qy 235 VRELINTNGARKNAFKILVITDGEKFGDPLGYEDVPIPADREGVIRYVIGVDAPRSE 294
Db 158 ATEVFRELGARPDATKVLIIITDGE--ATDSGNIDAAXD-----IIRVIIGIKHFQTK 210

```

```

Qy 295 KSRQELNTIASKPPRDHYPOVNNFEALKTIONLREKIPALEGTOTGSSSSFEHMSQSG 354
Db 211 ESEQLHKPASPASEFVKILDTFKLDFLTELQKIIYIEGTSKQDITFNMELUSSSG 270
Qy 355 FGAATISNGPLSLTVSGSDVWAGGVP-LYTSKEKSTFINNMTRVDSMDNDAYLFAAA-IIL 412
Db 271 ISADLSRGHAAVVGAGVAGXWAGGFLDLKADLQDDTFIGNELPLTPPEVRAGLYGVTVTLPS 330
Qy 413 RNRVQSLVLCAPRYQHIGLVAMFR--QNTGMFESNANVKGTOIGAYFGASLCSVDVDSNG 470
Db 331 ROKTSLASGAPRYQHMGREVLFPQBPQGGHWSQVTHGTQIGSYFGELCGVDVDDQG 390
Qy 471 STDVLIGAPHYEQTRGQVSVCLPFCQRARWQDAV--LYGQGGQPMGFGAALTYL 528
Db 391 ETELLLIIGAPFYGQRGGRVFYI-----QRQLGFEFVSVBLQDGPYPLGRFGAITAL 445
Qy 529 GDVNGDKLTDVAIGAPGBEDNRGAVLFGHTSGSGISPSHSORIASGSKLSPLOVFGQSL 588
Db 446 TIDNGDGLVDVAVGAPLEE--QGAVYIFNGRHG-GLSPQSORIEGTQVLSGIOMFGRSI 502
Qy 589 SGGQDLTMDGLVLTGAGHVLRLRSQPLVAVKAIMBPNPREVARNVPECDQVQV-KGK 647
Db 503 HGVKDLGEGDLADVAAGAESQMVILSSRPVDMVTLMSFPAEIPVHEVECSYSTSKMK 562
Qy 648 EAGEVRVCLHVOKSTRDLRREGQIOSVVTYDIALDSGRPHSRVAVNETKNTSTRQOVLG 707
Db 563 EGVNITICFQI-KSLIPQF-QGRLVANLTYTLQLDGHRTREGRGLPFGGRHRLRNIAVT- 619
Qy 708 LTOTCETLKLQLPNCIEDPVSPVILRLNPSL---VGTPLSAPCN-----LRPLAEDAQ 758
Db 620 TSMSCDTSDFPFCVQDLISPINVSLNFSLWEBSGTFRDQDQAGKDIPILAPSLHSETW 679
Qy 759 RLFTALFPFKCKGNNDNICODDLSTFTSFMSLDCLVWGSPREFNVTVTVRNDEGDSYRTQ 818
Db 680 EI-----PFKNCGEDKCEANLRVSFSPARSALRLTAFASLSVELSLNLEEDAYWYQ 734
Qy 819 VTFFFPDLISYRKVSTLQNRORSORSLACES--ASSTEVSGALKSTSCSINHPIPPENS 876
Db 735 LDLPFPGLSFRKVMEL---KPHSQIPVSCCELPPESRLLSRAL---SCNVSPIPKAGH 788
Qy 877 EYTFNITFDVDSKASLGNKLLKANVTSENN----MPRTNTEFOLELPVKVAVYVMTS 932
Db 789 SVALQMFNTLVNWSGSDSVLHANVTCNEDSDLEEDNSATTI---IPILPINILIQD 845
Qy 933 HGVSTKYNFTASENTSRVMOHOYOV---SNLQSSLP-LSLAVFLVPLVRLNQTWDRPQ 988
Db 846 QEDSTLYVSFTPKGPKIHQVRKMYOVRQPSIHDHNIPTLEAVGVGPVPPSGPITHQWS 905
Qy 989 VTFSNLSSTCHTK--ERLPSHSD--FLAELRKAFVNVNCSIAVCQRIQCDIPFFGIQEEF 1044
Db 906 VQMBEPV--PCHVEDLERLPDAAEPCLPGALPRCPVW-----FRQEI 945
Qy 1045 NATLKGNSLFDWYIKTSHNHLIIYSTHILFNDVSFTLLPGQCAFVRSQETETKVEPFPV 1104
Db 946 LVQVITGLVELGVEIAS-SMFSCLSSLSISFNSSKHFLYGSNASL-AQVWKKVDVVEYK 1003
Qy 1105 NPLPIVGVSSVGGILLALITAAALYKLGFFPKQYKDMASEG-GPPGAPB 1152
Db 1004 QMLYLYVLSGIGLLILLIIFVLYKVGFFKXENLKEKWEAGEGVNGIP 1052

RESULT 12
Q8HZVO PRELIMINARY; PRT; 927 AA.
AC Q8HZVO;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Lymphocyte function-associated antigen 1 (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

```

[illegible]

D5	63	NRNTAALYRCALSI - SPDEIACQ - PVEHICMPKRGYQGVTLVGNHGVLCVQVQARKFR	120
QY	114	SENTYVKGCLFLPGSNLRQOPKPPALRG -	144
D5	121	SLNSELTGACSLTLPNLDLQAQAYFSDLEGFLDPGARVDSDGYCRSGKGSTGEEKSARR	180
QY	145	-----POBDSIAFLDGGSSIIPEHPRMKFVSTVMEQL--KSKTLPLSLMOYS	193
D5	181	RRVREDEDEDEGTEIAVLVDGSGSTPSDFQAKNPFISTMMNFYEFKFCFNCFPALVQVG	240
QY	194	BEFRIHTFKFQNNPNRSLVKPIITQLGRTHATGIRKVRBELFNITNGARKNAFKIL	253
D5	241	AVIQTDFDLQESRDINASLAKVQSIQVKEVTKTASAMQHVLDNIPIPSGRGSKALKVM	300
QY	254	VWITDEKFGDPLGYEDVPEADREGVIRVIGVGDAFSEKSRQELNTIASKPRDHPV	313
D5	301	VWITDGDIFGDLNLTIVNSPKQGVVPAFGDAPFNNNTYRELKLIASDPKEAHTF	360
QY	314	QVNNFALKTIQOLREKIFAPEGTQTGSSSFEHMSQEGFSAATISNGP--LLSTVGSY	372
D5	361	KVTNYGALDGLLSKLOORIYHMEGT--VGDALQVLAQTGFSAQILDKGVLGTGVGAP	417
QY	373	DWAGGVFLY--TSKEKSTFINMT--RVDS--DMNDAYLGYAAAILLRNVQSLVLCAPRYQHI	429
D5	418	NWGGGALLYSTQNGRCFLNQTKADESRVTQVSYLSIAVLHKAHGISYVACAPHHKL	477
QY	430	GLVAMFRONTGMWESNA--NVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQT	486
D5	478	GAVFELKXEDR--EEDAFVRIISGEQMSYFGSVLCFVDINDGTTDFLVAAPFHIRG	535
QY	487	RGQSVVCLPQROARWQDDAVLYGEQOPKGRFGAALTVLGDVNGDKLTVAAGAP--	544
D5	536	EEGRVTVYQVPE--QDASFSLAHTLSHPGLTNSRFGFNMAAVGDTNODKFTDVAIGAPLE	594
QY	545	-----GEEDNRGAVYLPHTGSGSISPSHSORIASKLSPLQVFGQSLSGGODLTMDGLV	600
D5	595	GFAGDGASGVYIYNGHSG--GLYDSPQOIPASSVAGSLHYFGNSVSGGLDFNGDGLA	653
QY	601	DLTVGAGHVLRLRSQPVLRVAKIMFNPREVARNVFNQDVQVKGKAGEVRVCLHVOK	660
D5	654	DIITVGRDSAVVLRSPVVDLTYSMTFTP-----DALPVPVIGKM--DYNLCFEVDS	703
QY	661	S---TRDLREGIQSVVTVDLALDSGRPHSRVAFNETKNSTRQTOVLGLTQC-----	712
D5	704	SVVASEPGLREMFINTVDVVDV-----TKQORQLCEDSSGSCQCLURKWN	748
QY	713	-----ETLKLQPLNCIEDPVSPVILRLNFSLVGTPLSAPGNLR-----PVLAED	756
D5	749	GGSPFLCEHFWLSTEEI-----CEDCFSNITIKVYE-----FQTSGGRRDVPNPTL--	797
QY	757	AOELFTALP--PREKCGNDNTCODDLSITFSPMSLDCLVVGCPREFNVTIVTRDGEDS	814
D5	798	HYKEPSAIFQLPFEKDKKRVPCIAEIQUTTN--ISQELVGVKTEVTWNISLITNSGEDS	856
QY	815	YRTOVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEVSGALKSTCSINHPIPPE	874
D5	857	YMINMALNYPRNLQPKI-----QKPVSPDVQCDPKPV--ASVLVNNCKIGHPII--K	906
QY	875	NSEVTNITFDVDSKASIGNKLLKANVTSENNMPTNKTETQLELPVKYAVYVMTSHG	934
D5	907	RSSVNVSVTWQLEESVFPNRTADITVITNSNEKSLARETR--SLOFPHAFIAYLSR--	961
QY	935	VSTKYLNFTASENTSRVMQHOVQVSNLQORSILPSLVPLVPVRLNCTVWDRPQVTFPSN	994
D5	962	PSWMTYN--TSQSPSDHKBEFFNVHGENLFGAVFQIQCVPKIQDF-----QIVRVKN	1013
QY	995	LST-----CHTKERLPSHSDFLAEURKAPVNVCSIAVCORIQCDDIPPGIOEEFNATLK	1049
D5	1014	LTKTDHTECTQSQBACGSDPVQVHKWHSVVCAI-----TSNK	1053
QY	1050	GMLSPDWYIKTSHNHLIIVSTA-----ELFNDSVFTLLPGQGVPRVQOTETKVEPF--	1101
D5	1054	ENVTVAAEISVGHTKQLLDVSELPIILGELISFNKSLYEGLNAE-----NHETKITVFL	1107

QY	1102	--EVNPLPLIVGSSVGLLLALITAAALYKLGPFKRYKDMWSE	1144
D5	1108	KEEETRSPLIIGSSIGLLVLLVWIIALLFRGPFKRYQQLNLE	1152
RESULT 14			
ID	O88341	PRELIMINARY;	PRT; 1167 AA.
AC	O88341;		
DT	01-NOV-1998 (TREMELrel. 08, Created)		
DT	01-NOV-1998 (TREMELrel. 08, Last sequence update)		
DT	01-NOV-2003 (TREMELrel. 24, Last annotation update)		
DE	Integrin alpha E2 (Fragment).		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
OX	NCBI_TaxID=10116;		
RP	[1]		
RE	SEQUENCE FROM N.A.		
RX	MEDLINE=98056820; PubMed=9394838;		
EA	Brenan M., Rees D.J.G.		
RT	"Sequence analysis of rat integrin alpha E1 and alpha E2 subunits:		
RT	tissue expression reveals phenotypic similarities between		
RT	intraepithelial lymphocytes and dendritic cells in lymph."		
RL	Eur. J. Immunol. 27:3070-3079(1997).		
DR	EMBL; AF020046; AAC23663.1; --		
DR	HSSP; P11215; 1BHQ.		
DR	GO; GO:0008305; C:integrin complex; IEA.		
DR	GO; GO:0004895; F:cell adhesion receptor activity; IEA.		
DR	GO; GO:0007160; P:cell-matrix adhesion; IEA.		
DR	InterPro; IPR000413; Integrin_alpha.		
DR	InterPro; IPR002035; VWF_A.		
DR	Pfam; PF01839; FG-GAP; 3.		
DR	Pfam; PF00357; Integrin_A; 1.		
DR	Pfam; PF00092; vwa; 1.		
DR	PRINTS; PR01185; INTEGRINA.		
DR	PRINTS; PR00453; VWFADOMAIN.		
DR	SMART; SM00191; Int_alpha; 4.		
DR	SMART; SM00327; VWA; 1.		
DR	PROSITE; PS00242; INTEGRIN_ALPHA; 1.		
DR	PROSITE; PS02334; VWFA; 1.		
FT	NON TER		
SQ	SEQUENCE 1167 AA; 128593 MW; E3PED7E18B5CBEC CRC64;		
Query Match	18.9%; Score 1125.5; DB 11; Length 1167;		
Best Local Similarity	27.8%; Pred. No. 1.6e-72;		
Matches 352;	Conservative 214; Mismatches 456; Indels 243; Gaps 41;		

QY	5	VLLLTALFLCHGNLDTENAMTFQENARGFGQSVVQLQGSRVVWGAQOEIVAAQNGSLY	64
D5	6	LLCVASLKLPLGAFNVVDVDSW-----VTALQ-----FGAPSVL-----NSLKH	43
QY	65	QCDYSTGSCPEIRLQVPVEAVNM--SLGLSLAATSPPLACGPTVH-----	110
D5	44	QDPSNNQTC-----LLVARRSSNRNSAVLYQCATSIIPEIGCQVPEHILMPKRYQGVTL	99
QY	111	-----OTCSNTYVYKGLCFLFGSNLRQOPKPPALRG-----	143
D5	100	VRNHNGLVLCIQVQSRKPRSLNSLTGACSLTNLDLQAOAYFSDLEGVLDGASVNSG	159
QY	144	-----CPQED-----SDIAFLDGGSSIIPEHPRMKFVS	174
D5	160	DYGSGKGGSTGEETKSARLRQAVBEDEBEACTEIAVLVDGSGSTPSDFQAKDPIS	219
QY	175	TVMEQL--KSKTLPLSLMOYSEEPRIHFTKPFQNNPNRSLVKPIITQLGRTHATGIR	232
D5	220	TMENFTYKCEPCFALVQVGGVITQTEFDLLSDSDINASLAKVQSIQVKEVTKTASAMQ	279
QY	233	KVRELFNITNGARKNAFKILVWITDGEKFDPLGYEDVPEADREGVIRVYVIGVDAPR	292
D5	280	HVLONIFIPSGRSKKALKVMVLTLDGDIPEFDPLNLTVTIIVSSSKMGVWRVPAIGVNAPE	339

```

QY 293 SEKSRQELNTIASKPPRHVFOVNFALKTIONOLREKIPAEIGTQTGSSSSPFHEMSQ 352
DB 340 NNTVRELKLIASDPKAAHTFKVYNSALDGLSKLQREIIMHEGT---VGDTLQYQIAQ 396
QY 353 EGFSAITNSGP-LLSTVGSYDAGGVFLY-TSKEKSTPNNTRVDSNDNA---YLGYA 407
DB 397 TGFSAQILDKGQVLLGTGAFNWSGALLYNTQNGRFLNQT-AKEDFRAAQSYLGYS 455
QY 408 AAILRNVRQSLVLGAPRVOHIGLVAMPFRONTGMWESNANVGTGIGAYFGASLCSVDVD 467
DB 456 VAAIHKAGVSVAGAPRHKLRGANPELOKEDGETFMRIBGEQMSYFGSVLCPVDIN 515
QY 468 SNGSTDLVLIGAPHYYEOTRGGOVSVCPLEPRQORARWQCDVAVLYGEOGQFNGRFGAALTU 527
DB 516 MDGITDFILVAAPFYHIREGGRVYIRV-HEQDAPFSLVYTLGYPGLTSSRFQFANAA 574
QY 528 LGDVNGDKLTVAICAP-----GREDNEGAVYLPHTGTSGSISPSHSORAGSKLSPRL 581
DB 575 VGDINQDKFTDVAIGAPLEGFAGDAGSYGVIYINGHSQ-GHASPQOQIRASSVALGL 633
QY 582 QYFGOSLGGQDLTMDGLVDLTGAGQGHVLLRSQFVLKRXVKAIMBPNPREVARNVFECDN 641
DB 634 YIFGMSVSGLDGDFSGDDLADITVGSQDVAVWLRSPVVDLTVSMFTPDALP----- 685
QY 642 QYVKGKEAGEVRVCLHVQXS---TRDLREGQIQSVVTVYDLDLDSGRHSRAVFNETGNS 698
DB 686 --MAFKDMVDVCLCFKVDSSAVPSEBGLRGMENFTVDVDV-----TKQK 728
QY 699 TRRQ-----TQVLGLTQCTCLKLQLENCIEDPVPISVLRLNLSL 738
DB 729 QRLQCADRSCCCLMKWSSGSLCEHFGLI STEBEL-----CEDDCFSNITIKVSEF 782
QY 739 VGTPLSAPGNLAPVLAARDQRLFTALP--PFRKNGDNICQDDLSITPFSMSLCLLVG 796
DB 783 -OTSEERRNHPIL--DHYKEPSAIFQLPYEKCKXKVFCAIEIQLTTAISQDP-LVVG 838
QY 797 GPREFNVTYVNDGEDSYRTQVTFPPLDLSYRKVSTLQNKORSQSRWKLACESASSTEV 856
DB 839 IYKEVTNLSLTNSGSDSWTNALNYPNQLQFKI-----QKPLSPDIQCDPKPV-- 890
QY 857 SGALKSTCSINHPIPPENSEVTNFTPDVDKASISGNKLLKANVTSENMPRNKTEF 916
DB 891 -ASVLVNMCKIHPIL-KRSSVNSVTVQLEESIFPNRTADITVITSNNEKSLARETH- 947
QY 917 QLELPVKYAVYVMTSHGVSSTYLNFTASNTSRVMOHOYOVSNLQORSLPISLFLVPV 976
DB 948 --SLQPRHAFIAVLSR--FSVMTMN--TSQSSDHKEFFNVHGENHFGAVQLQICVPI 1001
QY 977 RLNQTVINDRPQVTFSENLSST-CHTKERLPSHSDFLAELRKAPVWNCVSIYVQRIQCDI 1035
DB 1002 TERDL-----QIIRVKLTKTQAHT-----ECTSQSEPTCGSDPVQNV---- 1039
QY 1036 PFGIOEENATL-----KGNLSFDWYIKTSHNHLI--VSTABIL-----FNDVSFTLL 1083
DB 1040 -----EHWHSVCAITSNKENTYAAEISMGHTKQLLRDISELQILGSEISFNKSLYEL 1093
QY 1084 PQGAFVRSQTEKVEPFF-----EVPNPLFIVGSSVGGILLALALITAAALYKLGFKRQYK 1139
DB 1094 NAE-----NHRKIIVIFLKEKEKPSLPIIGSSIGGLLVLVVLIALLFKCGFFKRYK 1147
QY 1140 DMMSSE 1144
DB 1148 QLNLE 1152

```

RESULT 15

O42094 PRELIMINARY; PRT; 1171 AA.

ID O42094

AC O42094;

DT 01-JAN-1998 (TEMBLrel. 05, Created)

DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)

DT 01-JUN-2003 (TEMBLrel. 24, Last annotation update)

DE ALPHA1 integrin.

```

OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus
OX NCBI_TaxID=9031;
RN [1]
RC SEQUENCE FROM N.A.
RC TISSUE=Gizzard;
RX MEDLINE=97476270; PubMed=9334246;
RA Obata H., Hayashi K., Nishida W., Momiyama T., Uchida A., Ochi T.,
RA Sobue K.;
RT "Sarcooth muscle cell phenotype-dependent transcriptional regulation of
RT the alpha1 integrin gene.";
RL J. Biol. Chem. 272:26643-26651(1997).
DR EMBL; AB000470; BAA23160.1; -
DR EMBL; AB000471; BAA23161.1; -
DR PIR; A55348; A55348.
DR HSSP; P17301; IAOX.
DR GO; GO:000305; C:integrin complex; IEA.
DR GO; GO:0004895; P:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR GO; GO:0007229; P:integrin-mediated signaling pathway; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWP_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PRO1185; INTEGRINA.
DR PRINTS; PRO0453; VWFADOMAIN.
DR SMART; SM00191; Int alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS50234; VWFA; 1.
DR Integrin.
KW SEQUENCE 1171 AA; 130228 MW; B505A4D65F09736E CRC64;
QY 1 MALRVLLLTAL-TLCHKFNLDTEAMTPQENARG-EGOSVVQL---QGSRVVVGAPQIV 55
DB 1 MAASVYMLTALFHLTGASNVVDVKNMTPSGPLEDMFGITVOOYENEGKNVLGSPLVQG 60
QY 56 AANQSGLSYQCDYSTGSGCEP-IRLQVP-----VEAV--NMSLGLSAAATSPQLLAC 105
DB 61 PEKRTGVDYKCPVGRDSSQPCIKLNPDATSPVNVMEVKENMTLGTTL-VTNPKGFLAC 119
QY 106 GPTVHTQTSNTYTKGLCFLEGSNLRQOPQKPPALRCQPOEDSDIALIDGSGIIPHD 165
DB 120 GPLYAYKGRLLHYTGVCSNVSSSTPETVKAVAP-SVQEC-KTQDDIVIVLDGNSIYP-- 175
QY 166 FRRMKPEFVSTVWELK--KSKTLPSLMQYSBEFRIHFTFKBFQNNPNRSLVKPTQLLG 223
DB 176 MESVTAFLNSLLRNMDIGPQTQVGIQYGVTVVHEFYLNTYSTTEEVDAALRIROGG 235
QY 224 -RTHFATGIRKVVRELNFNTNGARKNAFKILVITDGEKFGDPLGYDVIPEARDEGVIR 282
DB 236 TQTMALGIDTAREAPTEAHGARRGVOKVMVIVTDGESH-DNYRLQEVLDKCEDENQR 294
QY 283 YVIGVGDAF-----RSEKSRQBLNTIASKPPDRDHFVQVNFPEALKTIONQLREKIFAIEG 337
DB 295 FAIALGSIYSGNLSSTFKFVEIKSIASKPTEKFFNVSDBLALVTIVALEGERIFALEA 354
QY 338 TQTGSSSFHEHMSQBFSAAITNGPLLLSTVGSYDAGGVFLYTSKEKSTPNNTRVDS 397
DB 355 ITDQQAASFEMEMSQAGFSAHYSDQWVNLGAVGYDMNGTVMVVKDSISIPSNDFRDR 414
QY 398 DMND-----AVLGYAAAAILLRNVQSLVGLGAPRYOHIGLVAMPFRONTGMWESNANVKGQ 452
DB 415 HSEKTEPLAALYGYTVNSALTPGGVLYTAGQPRNHTQVLIYKMEGSEVQLQELKSEQ 474
QY 453 IGAYFGASLCSVDVDSNGSTDLVLIGAPHYY--EQTRGGQVSVCPFLPRG-----QRA 502

```

Query Match 17.9%; Score 1068.5; DB 13; Length 1171;
 Best Local Similarity 28.4%; Pred. No. 2.3e-68;
 Matches 349; Conservative 215; Mismatches 521; Indels 145; Gaps 47;

Search completed: June 7, 2004, 17:16:09
Job time : 48.0446 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 16:54:39 ; Search time 50.1568 Seconds
(without alignments)
6495.175 Million cell updates/sec

Title: US-09-902-481B-1

Perfect score: 5956
Sequence: 1 MALRVLLTALTLCGHFNLD.....FRQYKDMSEGGPGAEPPQ 1153

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_25Jan04: *
1: Geneseqp1980s: *
2: Geneseqp1990s: *
3: Geneseqp2000s: *
4: Geneseqp2001s: *
5: Geneseqp2002s: *
6: Geneseqp2003as: *
7: Geneseqp2003bs: *
8: Geneseqp2004s: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	5956	100.0	1153	2	Aaw65090 Human Bet
2	5956	100.0	1153	3	Aab07360 Human CDI
3	5956	100.0	1153	5	Aau80252 Human Int
4	5956	100.0	1153	5	Abg61469 Human Bet
5	5956	100.0	1153	5	Aaol1428 Integrin
6	5956	100.0	1153	7	Add25615 Binding d
7	5946	99.8	1153	2	Aar04136 Alpha sub
8	3494	58.7	1163	2	Aar07120 p150.95 a
9	3480	58.4	1163	2	Aaw65091 Human Bet
10	3480	58.4	1163	3	Aab07361 Human CDI
11	3480	58.4	1163	5	Abg61470 Human Bet
12	3478	58.4	1163	6	Abu07406 Protein d
13	3455	58.0	1161	2	Aar78166 Human Bet
14	3455	58.0	1161	2	Aaw23049 Human Bet
15	3455	58.0	1161	2	Aaw57491 Human Bet
16	3455	58.0	1161	2	Aaw65089 Human Bet
17	3455	58.0	1161	2	Aaw72825 Human Bet
18	3455	58.0	1161	2	Aaw73342 Human alp
19	3455	58.0	1161	3	Aab07359 Human alp
20	3455	58.0	1161	5	Abg61468 Human Bet
21	3439.5	57.7	1161	2	Aaw23064 Human Bet
22	3439.5	57.7	1161	2	Aaw65306 Human Bet
23	3439.5	57.7	1161	2	Aaw72837 Human alp
24	3439.5	57.7	1161	2	Aaw73343 Human alp
25	3439.5	57.7	1161	3	Aab07376 Human alp

26	3439.5	57.7	1161	5	ABG61485	Abg61485 Human Bet
27	3272.5	54.9	1161	2	AAR78169	Aar78169 Rat alpha
28	3270.5	54.9	1161	2	AAW23062	Aaw23062 Rat beta
29	3270.5	54.9	1161	2	AAW60004	Aaw60004 Rat alpha
30	3270.5	54.9	1161	2	AAW72824	Aaw72824 Rat alpha
31	3270.5	54.9	1161	3	AAAB07374	Aab07374 Rat alpha
32	3270.5	54.9	1161	5	ABG61483	ABG61483 Rat Beta2
33	3264	54.8	1161	2	AAW23061	Aaw23061 Mouse bet
34	3264	54.8	1161	2	AAW60003	Aaw60003 Mouse bet
35	3264	54.8	1161	2	AAW65103	Aaw65103 Mouse bet
36	3264	54.8	1161	2	AAW72836	Aaw72836 Mouse alp
37	3264	54.8	1161	2	AAW73347	Aaw73347 Mouse alp
38	3264	54.8	1161	3	AAAB07373	AAAB07373 Mouse alp
39	3264	54.8	1161	5	ABG61482	ABG61482 Mouse Bet
40	3263.5	54.8	1161	2	AAW65104	Aaw65104 Rat beta-
41	3263.5	54.8	1161	2	AAW73345	Aaw73345 Rat alpha
42	3260	54.7	1161	2	AAR78168	Aar78168 Mouse alp
43	3251.5	54.6	1151	2	AAW23059	Aaw23059 Rat beta
44	3251.5	54.6	1151	2	AAW60001	Aaw60001 Rat alpha
45	3251.5	54.6	1151	2	AAW65101	Aaw65101 Rat beta-

ALIGNMENTS

RESULT 1

AAW65090
ID AAW65090 standard; protein, 1153 AA.

XX AC AAW65090;

XX DT 28-SEP-1998 (first entry)

XX DE Human Beta-integrin CD11b subunit protein.

XX KW Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;
XX KW type-I diabetes; atherosclerosis; multiple sclerosis; aschma;
XX KW lung inflammation; acute respiratory distress syndrome; CD11b subunit;
XX KW rheumatoid arthritis.

XX OS Homo sapiens.

XX FN US5728533-A.

XX PD 17-MAR-1998.

XX PF 07-JUN-1995; 95US-00485618.

XX PR 23-DEC-1993; 93US-00173497.

XX PR 05-AUG-1994; 94US-00286889.

XX PR 21-DEC-1994; 94US-00362652.

XX PA (ICOS-) ICOS CORP.

XX PI Van Der Vieren M, Gallatin WM;

XX XX WPI; 1998-206565/18.

XX DR Screening assay for modulators of integrin binding - using immobilised or

XX CC labelled alpha-d polypeptide, useful for, e.g. treating type-I diabetes.

XX PS Example 5; Fig 1A-D; 106pp; English.

XX CC This sequence represents a human beta-integrin CD11b subunit which is
XX CC used to describe a method for identifying compounds that modulate the
XX CC interaction of the beta-integrin alpha-d subunit with a binding partner
XX CC of alpha-d which involves contacting an alpha-d polypeptide with an alpha
XX CC -d binding partner, one of which is immobilised and the other of which is
XX CC labelled, in the presence of a test compound, and determining if the
XX CC compound affects binding between the alpha-d polypeptide and alpha-d
XX CC binding partner, where the alpha-d polypeptide is alpha-d or its fragment
XX CC comprising the cytoplasmic, transmembrane or extracellular domain of
XX CC alpha-d. Compounds that modulate alpha-d binding could be used to treat

CC diseases such as type-I diabetes, atherosclerosis, multiple sclerosis,
CC asthma, psoriasis, lung inflammation, acute respiratory distress syndrome
CC and rheumatoid arthritis
XX
SQ Sequence 1153 AA;

Query Match 100.0%; Score 5956; DB 2; Length 1153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRVLLTALTLCCHFNLDENAMTFQENARGFGQSVVQLQGRVVVGAPQBIIVAANOR 60
DB 1 MALRVLLTALTLCCHFNLDENAMTFQENARGFGQSVVQLQGRVVVGAPQBIIVAANOR 60

QY 61 GSIYQCDYSGSCSEPTRLQVPEAVNMSLGLSLAATSPQILLACQPTVHQCSEMYTK 120
DB 61 GSIYQCDYSGSCSEPTRLQVPEAVNMSLGLSLAATSPQILLACQPTVHQCSEMYTK 120

QY 121 GLCFLFGSNLRQOQPFPPALRCPCQEDSDIAFLIDSGSIIPHDFRMKEFVSTWEOI 180
DB 121 GLCFLFGSNLRQOQPFPPALRCPCQEDSDIAFLIDSGSIIPHDFRMKEFVSTWEOI 180

QY 181 KKSRTLPSLMQYSEEBRIHPTFKFQNNPNPSLVKPIITQLLGRTHATGIRKVVRELPN 240
DB 181 KKSRTLPSLMQYSEEBRIHPTFKFQNNPNPSLVKPIITQLLGRTHATGIRKVVRELPN 240

QY 241 ITNGARKNAFKILVITDGEKCDPLGYEDVPEADREGVIRVVGVPAPFSEKSRQEL 300
DB 241 ITNGARKNAFKILVITDGEKCDPLGYEDVPEADREGVIRVVGVPAPFSEKSRQEL 300

QY 301 NTIASKPPRDHVPQVNNFEALKTIONLREKIPAEICTGTGSSSSPEHMSDEGFSAAIT 360
DB 301 NTIASKPPRDHVPQVNNFEALKTIONLREKIPAEICTGTGSSSSPEHMSDEGFSAAIT 360

QY 361 SNGPLASTVGSYDMAGVPLTYSKEKSTFTNTRVDSMDNDAYLGVAARAILLRNVQSLV 420
DB 361 SNGPLASTVGSYDMAGVPLTYSKEKSTFTNTRVDSMDNDAYLGVAARAILLRNVQSLV 420

QY 421 LGAPRYQHIGLVAMFRQNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAP 480
DB 421 LGAPRYQHIGLVAMFRQNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAP 480

QY 481 HYVEOTRGQGVSVCPILPRGQARWQCDVLYGEGQGPWGRFGAALFVLGDVNDKDLTDVA 540
DB 481 HYVEOTRGQGVSVCPILPRGQARWQCDVLYGEGQGPWGRFGAALFVLGDVNDKDLTDVA 540

QY 541 IGAPGEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQVFGQSLSGGQDLTMDGLV 600
DB 541 IGAPGEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQVFGQSLSGGQDLTMDGLV 600

QY 601 DLTVGAQCHVILLRSQVPLRVKAIMFNPREVARNYFPCNDQVVKKEAGFVRVCLHVOK 660
DB 601 DLTVGAQCHVILLRSQVPLRVKAIMFNPREVARNYFPCNDQVVKKEAGFVRVCLHVOK 660

QY 661 STDRRLREGQIQSVVYTDALDGRPHSRVAFNETKNSTRROTQVILGLTQTCETLKLQLP 720
DB 661 STDRRLREGQIQSVVYTDALDGRPHSRVAFNETKNSTRROTQVILGLTQTCETLKLQLP 720

QY 721 NCTEDVPVTLRLNLSVLTPLSAFNGNLRLVLAEDAQRLLFTALPPFNKQGNNDICQD 780
DB 721 NCTEDVPVTLRLNLSVLTPLSAFNGNLRLVLAEDAQRLLFTALPPFNKQGNNDICQD 780

QY 781 LSTTFMSLUDCLVWGSPREFNTVTRVNDGDSYRTQVTFPPDLDSYRKYSTLQNGRS 840
DB 781 LSTTFMSLUDCLVWGSPREFNTVTRVNDGDSYRTQVTFPPDLDSYRKYSTLQNGRS 840

QY 841 QRSWLACSSASTSVSGALKSTSCINHPIPPENSEVTFNITFDVDSKASLGNKLLKA 900
DB 841 QRSWLACSSASTSVSGALKSTSCINHPIPPENSEVTFNITFDVDSKASLGNKLLKA 900

QY 901 NVTSENNMPRTNKTBEQLELPVKYAVYVMTVTSKVTLNFTASENTSRVWQHOYQVSN 960
DB 901 NVTSENNMPRTNKTBEQLELPVKYAVYVMTVTSKVTLNFTASENTSRVWQHOYQVSN 960

QY 961 LGQRLSPISLVFLVPLVRLNQTVWDRPQVTFSENLSSTCHTKERLPSSHDFLAELRKAPV 1020
DB 961 LGQRLSPISLVFLVPLVRLNQTVWDRPQVTFSENLSSTCHTKERLPSSHDFLAELRKAPV 1020

QY 1021 VNCISIAVCORICQDIPPGIOEENATLKGNSLFDWYIKTSHNHLIIVSTAELLFNDSPV 1080
DB 1021 VNCISIAVCORICQDIPPGIOEENATLKGNSLFDWYIKTSHNHLIIVSTAELLFNDSPV 1080

QY 1081 TLLPQCGAFVRSQTEKVEPEFVNPPLIIVGSSVGGLLALITAALYKLGFFPRQYKD 1140
DB 1081 TLLPQCGAFVRSQTEKVEPEFVNPPLIIVGSSVGGLLALITAALYKLGFFPRQYKD 1140

QY 1141 MMSBGGPPGAEPPQ 1153
DB 1141 MMSBGGPPGAEPPQ 1153

RESULT 2
AAB07360
ID AAB07360 standard; protein, 1153 AA.
XX
AC AAB07360;
XX
DT 17-JAN-2001 (first entry)
XX
Human CD11b protein sequence.
DE
XX
Human; macrophage infiltration inhibition; alpha_d integrin;
KW leukocyte integrin; Leu-CAM; leukointegrin; immune response;
KW inflammation; leukocyte adhesion deficiency; LAD; Type I diabetes;
KW atherosclerosis; multiple sclerosis; asthma; psoriasis; beta2 integrin;
KW lung inflammation; acute respiratory distress syndrome; Crohn's disease;
KW rheumatoid arthritis; central nervous system injury; CD11b.
XX
OS Homo sapiens.
XX
FN WO200029446-A1.
XX
PD 25-MAY-2000.
XX
PF 16-NOV-1999; 99WO-US027139.
XX
PR 16-NOV-1998; 98US-00193043.
PR 08-JUL-1999; 99US-00350259.
XX
PA (ICOS-) ICOS CORP.
XX
PI Gallatin MW, Van Der Vieren M;
XX
DR WPI; 2000-387751/33.
XX
PT Use of novel anti-alpha integrin d monoclonal antibodies to inhibit
PT macrophage infiltration and reduce inflammation at central nervous system
PT injury sites.
XX
PS Example 5; Fig 1; 270pp; English.
XX

Integrins are a class of membrane-associated molecules that participate in cellular adhesion. Integrins are made up of an alpha subunit and a beta subunit. One class of human integrins are restricted to expression in white blood cells and have a common beta2 subunit: the leukocyte integrins. Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins have an important role in immune and inflammatory responses. The present protein sequence is the human integrin alpha subunit CD11b. This sequence was used in an alignment to identify a novel beta2 integrin alpha subunit: alpha_d (AAA60014 and AAB07359). The present sequence has approximately 50% identity to the protein sequence of alpha_d. The Alpha_d gene and protein may be useful in therapy for diseases linked to aberrant alpha_d function e.g. Type I diabetes, atherosclerosis, multiple sclerosis, asthma, psoriasis, lung inflammation, acute respiratory distress syndrome, rheumatoid arthritis and leukocyte adhesion deficiency (LAD). In addition, anti-alpha_d monoclonal antibodies may be used in the

CC inhibition of macrophage infiltration at the site of a central nervous
CC system injury. The monoclonal antibodies can also be used to detect and
CC diagnose Crohn's disease
XX
SQ Sequence 1153 AA;

Query Match 100.0%; Score 5956; DB 3; Length 1153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRVLLLTALTLCHGFNDLTENAMTFQENARFGQSVVQLOGSRVVVCGAPDIIIVAAQR 60
DB 1 MALRVLLLTALTLCHGFNDLTENAMTFQENARFGQSVVQLOGSRVVVCGAPDIIIVAAQR 60

QY 61 GSIYQCDYSGCEPIRLQVPVEAVNMSLGLSLAATSPQLLACGPTVHQTCSENTYVK 120
DB 61 GSIYQCDYSGCEPIRLQVPVEAVNMSLGLSLAATSPQLLACGPTVHQTCSENTYVK 120

QY 121 GICFLFGSNLRQOPQPFALRCPOEDSDIAPLDGSGSIIPHPFRMKFVSTVMEQL 180
DB 121 GICFLFGSNLRQOPQPFALRCPOEDSDIAPLDGSGSIIPHPFRMKFVSTVMEQL 180

QY 181 KKSCTLFLSMQYSEBEIRIHTFKFQNNPNRSLVKPITQLLGRTHTATGIRKVVRELFN 240
DB 181 KKSCTLFLSMQYSEBEIRIHTFKFQNNPNRSLVKPITQLLGRTHTATGIRKVVRELFN 240

QY 241 ITNGARKNAFKILVITDGEKCDPLGYEDVTPADREGVIRYVIGVDAFSEKSRQBL 300
DB 241 ITNGARKNAFKILVITDGEKCDPLGYEDVTPADREGVIRYVIGVDAFSEKSRQBL 300

QY 301 NTIASPPRDHVFQVNFALKTIQOLREKIPAEIGTGTGSSSPFEHMSQEGPSAALT 360
DB 301 NTIASPPRDHVFQVNFALKTIQOLREKIPAEIGTGTGSSSPFEHMSQEGPSAALT 360

QY 361 SNGPLLSVGSYDAGVFLYTSKESFTINMTRVDSMDNDAYLGYAAAILRNRVQSLV 420
DB 361 SNGPLLSVGSYDAGVFLYTSKESFTINMTRVDSMDNDAYLGYAAAILRNRVQSLV 420

QY 421 IGAAPRYOHLGLVAMFQNTGMESNANVKGTOIGAYFGASLCSVDVDSNGSDLVILGAP 480
DB 421 IGAAPRYOHLGLVAMFQNTGMESNANVKGTOIGAYFGASLCSVDVDSNGSDLVILGAP 480

QY 481 HYTEQTRGGQSVCPFLRQARWQCDVLYGEGQGFQWGRFGAALTVLGDVNGDKLTDA 540
DB 481 HYTEQTRGGQSVCPFLRQARWQCDVLYGEGQGFQWGRFGAALTVLGDVNGDKLTDA 540

QY 541 IGAAGEDNRGAVYLFHGTGSGISPSHSORAGSLSPLOYFGSLSGGQDLTMDGLV 600
DB 541 IGAAGEDNRGAVYLFHGTGSGISPSHSORAGSLSPLOYFGSLSGGQDLTMDGLV 600

QY 601 DLTVGAQGHVLLRSQPVLRVKALMEFNPREVARNVFECNDQVVRKGEAGEVVRCLHVQK 660
DB 601 DLTVGAQGHVLLRSQPVLRVKALMEFNPREVARNVFECNDQVVRKGEAGEVVRCLHVQK 660

QY 661 STRDLRREGIQSVVYDIALDSGRPHSAVFNKTNSTRROTQVLGTOTCETIKQLP 720
DB 661 STRDLRREGIQSVVYDIALDSGRPHSAVFNKTNSTRROTQVLGTOTCETIKQLP 720

QY 721 NCTEDPVSPIVLRNAPSLVGTPLSAPGNLRPVLAEDAQELFTALPFPEKNCNDNICODD 780
DB 721 NCTEDPVSPIVLRNAPSLVGTPLSAPGNLRPVLAEDAQELFTALPFPEKNCNDNICODD 780

QY 781 LSITPSFMSLDCLVGGPREFNVTVTRNDGSDSYRTQVTFPDLDSYRKVSTLQNGRS 840
DB 781 LSITPSFMSLDCLVGGPREFNVTVTRNDGSDSYRTQVTFPDLDSYRKVSTLQNGRS 840

QY 841 QRSWRLACSASTEVSGALKSTSCSINHPIPEPENSEVTFNITPDVDSKASLGNKLLKA 900
DB 841 QRSWRLACSASTEVSGALKSTSCSINHPIPEPENSEVTFNITPDVDSKASLGNKLLKA 900

QY 901 NVTSENMPRTNKTPEQLELPVKYAVYVMTVSHGVSTKYLNFTASENTSRVWQHQYQVSN 960
DB 901 NVTSENMPRTNKTPEQLELPVKYAVYVMTVSHGVSTKYLNFTASENTSRVWQHQYQVSN 960

QY 961 LQORSPLISLVPLVPLVRLNQTVIWDROPVTFSENLSTCHTKERLPSHSDFLAELRKAPV 1020
DB 961 LQORSPLISLVPLVPLVRLNQTVIWDROPVTFSENLSTCHTKERLPSHSDFLAELRKAPV 1020

QY 1021 VNCIAVCQRIQCDIPFFGIOEFENATLKGNSLSDWYIKTSENHLLIIVSTABILFNDSVF 1080
DB 1021 VNCIAVCQRIQCDIPFFGIOEFENATLKGNSLSDWYIKTSENHLLIIVSTABILFNDSVF 1080

QY 1081 TLLPQCGAFVRSQETKVEPFEVNPFLPLIYGVSSVGGELLALITPAALYKLGFFPKROYKD 1140
DB 1081 TLLPQCGAFVRSQETKVEPFEVNPFLPLIYGVSSVGGELLALITPAALYKLGFFPKROYKD 1140

QY 1141 NMSEGGPPGAPQ 1153
DB 1141 NMSEGGPPGAPQ 1153

RESULT 3
AAU80252
ID AAU80252 standard; protein; 1153 AA.
XX
AC AAU80252;
DT 15-JUL-2002 (first entry)
XX Human integrin 1 alpha-M subunit protein.
DE Integrin; antiinflammatory; immunosuppression; nephritis; dermatitis;
KW inflammatory disease; autoimmune disorder; Crohn's disease;
KW human immunodeficiency virus; HIV; myocardial infarction;
KW Sjorgen's syndrome; rheumatoid arthritis.
XX Homo sapiens.
OS
XX Key Location/Qualifiers
FT Misc-difference 499..500 /note= "Encoded by GGG CAG AGG"
XX WO200218583-A2.
XX 07-MAR-2002.
XX 31-AUG-2001; 2001WO-US027227.
XX 01-SEP-2000; 2000US-0229700P.
XX (BLOO-) CENT BLOOD RES INC.
XX Springer TA, Shimoaka M, Lu C;
XX WPI; 2002-382964/41.
XX N-PSDB; ASKS0046.
XX Modified integrin-I or integrin I-like domain polypeptide useful as an
XX immunogen to produce antibodies specific to polypeptide, comprises a
XX disulfide bond such that polypeptide is stabilized in a desired
XX conformation.
XX Disclosure; Page 109-112; 112pp; English.
XX This invention relates to a modified integrin-I or integrin I-like domain
XX polypeptide comprising at least one disulfide bond so that the domain is
XX stabilised in a desired conformation. The polypeptide of the invention
XX may have antiinflammatory or immunosuppressive activities. The
XX polypeptides of the invention have an open conformation and are useful as
XX immunogens to produce antibodies that selectively bind to integrin I-
XX domain; and for identifying a modulator of integrin activity, or of
XX interaction of an integrin and a cognate ligand. The polypeptide of the
XX invention, or antibodies (preferably anti-LFA-1 antibody) is useful for
XX treating or preventing an integrin mediated disorder which is an
XX inflammatory or autoimmune disorder in a subject and for inhibiting the
XX binding of an integrin to a cognate ligand such as Crohn's disease,

CC nephritis; human immunodeficiency virus (HIV), myocardial infarction,
CC Sjoren's syndrome, rheumatoid arthritis, dermatitis. A therapeutic
CC composition comprising the peptide of the invention is useful for
CC treating an integrin mediated disorder in a subject. The polypeptides
CC and/or active or antigenic fragments are useful as reagents for diagnosis
CC of integrin-mediated disorders. The present sequence represents the human
CC integrin-1 alpha-M protein subunit used to generate the mutant
CC polypeptides of the invention
CC
SQ Sequence 1153 AA;

Query Match 100.0%; Score 5956; DB 5; Length 1153;
Best Local Similarity 100.0%; Pred. NO. 0;
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MALRVLLLTALTCHGNLDTENAMTQENARGFGQSVVQLQSGRVVVGAPQIIVANQR 60
Db 1 MALRVLLLTALTCHGNLDTENAMTQENARGFGQSVVQLQSGRVVVGAPQIIVANQR 60

Qy 61 GSLYQCDYSTGSCPEPIRLQVPEAVNMSGLSLAATSPPLLACGPTVHQTSEITYVK 120
Db 61 GSLYQCDYSTGSCPEPIRLQVPEAVNMSGLSLAATSPPLLACGPTVHQTSEITYVK 120

Qy 121 GLCLFPGSNLRQPKPEALRGCPQSDIAFLIDGSGSIIIPHDPRMKEFYSTVWBL 180
Db 121 GLCLFPGSNLRQPKPEALRGCPQSDIAFLIDGSGSIIIPHDPRMKEFYSTVWBL 180

Qy 181 KSKTLFSLMOYSEBFIHTFKEFQNNPRLSVKPIITQLGRTHTATGIRKVVRELEN 240
Db 181 KSKTLFSLMOYSEBFIHTFKEFQNNPRLSVKPIITQLGRTHTATGIRKVVRELEN 240

Qy 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVIPEADREGVIRVIVGVDAFRSEKSRQEL 300
Db 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVIPEADREGVIRVIVGVDAFRSEKSRQEL 300

Qy 301 NTIASKPRDHFVQNNPEALKTIONLREKIFAIEGTQTGSSSFEHMSQSGPSAAIT 360
Db 301 NTIASKPRDHFVQNNPEALKTIONLREKIFAIEGTQTGSSSFEHMSQSGPSAAIT 360

Qy 361 SNGPLLSVGSYDWAGVGLVLTSEKSTFINMTRVDSMDNDVILGAAAILLNRVQSLV 420
Db 361 SNGPLLSVGSYDWAGVGLVLTSEKSTFINMTRVDSMDNDVILGAAAILLNRVQSLV 420

Qy 421 LGAPRYQHI GLVAFRQNTGMSNANVKTQIGAYFGASLCSVDVDSNGSTDLVLIGAP 480
Db 421 LGAPRYQHI GLVAFRQNTGMSNANVKTQIGAYFGASLCSVDVDSNGSTDLVLIGAP 480

Qy 481 HYETRGQGVSVCPPLRGQARWQCDVLYGQGPWGFGAALTVLGVNCDKLTDA 540
Db 481 HYETRGQGVSVCPPLRGQARWQCDVLYGQGPWGFGAALTVLGVNCDKLTDA 540

Qy 541 IGAPGEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLV 600
Db 541 IGAPGEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLV 600

Qy 601 DLTVGAQHVLRLRSQVPLRVKAIMFNPEVARNVPECDQVVKKEAGEVTVCLHVOK 660
Db 601 DLTVGAQHVLRLRSQVPLRVKAIMFNPEVARNVPECDQVVKKEAGEVTVCLHVOK 660

Qy 661 STRDLREGQIQSVVYDLDLDSGRPHSRVAFNETKNSFRQTQVLTOTCTETKLQLP 720
Db 661 STRDLREGQIQSVVYDLDLDSGRPHSRVAFNETKNSFRQTQVLTOTCTETKLQLP 720

Qy 721 NCIEDVSPVLRNLFSLVGTPLSAFGLNRPVLAEDAQRLLFTALPFPEKNCNDNTCCDD 780
Db 721 NCIEDVSPVLRNLFSLVGTPLSAFGLNRPVLAEDAQRLLFTALPFPEKNCNDNTCCDD 780

Qy 781 LSITFSPMSLDCLVVGSPREFNTVTVRNDGDSYRQTQVTFPPLDLVYKVKSTLQNRS 840
Db 781 LSITFSPMSLDCLVVGSPREFNTVTVRNDGDSYRQTQVTFPPLDLVYKVKSTLQNRS 840

Qy 841 QRSWRLACESASTSVSGALKSTSCSINHPIPPENSEVTFNITPDVDVSKASLGNKLLKA 900
Db 841 QRSWRLACESASTSVSGALKSTSCSINHPIPPENSEVTFNITPDVDVSKASLGNKLLKA 900

Db 841 QRSWRLACESASTSVSGALKSTSCSINHPIPPENSEVTFNITPDVDVSKASLGNKLLKA 900
Qy 901 NVTSENMPRTWNTKTFQLELPVKYAVYVMTVSHGVSTKYLNFTASNTSRVMQHOYQVSN 960
Db 901 NVTSENMPRTWNTKTFQLELPVKYAVYVMTVSHGVSTKYLNFTASNTSRVMQHOYQVSN 960
Qy 961 LQORSILPISLVLVPLVRLNQTVWDRPQVTFPSENLSSTCHTKERLPSSHSDFLAELKAPV 1020
Db 961 LQORSILPISLVLVPLVRLNQTVWDRPQVTFPSENLSSTCHTKERLPSSHSDFLAELKAPV 1020

Qy 1021 VNCSTAVCORIOCDIPFFGIQEEFNATLKNLSPDWYIKTSHNHLIIVSTABILFNDSVF 1080
Db 1021 VNCSTAVCORIOCDIPFFGIQEEFNATLKNLSPDWYIKTSHNHLIIVSTABILFNDSVF 1080

Qy 1081 TLLPQCGAFVRSQSTETKVEPPEVNPPLPLIVGVSSVGGILLALITAAALYKLGFFKQYKD 1140
Db 1081 TLLPQCGAFVRSQSTETKVEPPEVNPPLPLIVGVSSVGGILLALITAAALYKLGFFKQYKD 1140

Qy 1141 MNSEGGPPGAEQ 1153
Db 1141 MNSEGGPPGAEQ 1153

RESULT 4
ABG61469
ID ABG61469 standard; protein; 1153 AA.
XX AC ABG61469;
XX DT 27-AUG-2002 (first entry)
XX Human Beta2 integrin alphaCD11b subunit.
XX DE Beta2 integrin; alphaD subunit; CD11c subunit; CD11b subunit; LAD;
XX KW leukocyte adhesion deficiency; inflammatory response; diabetes;
XX KW multiple sclerosis; arthritis; graft atherosclerosis; neuroprotective;
XX KW inflammatory bowel disease; Crohn's disease; ulcerative colitis;
XX KW immune complex alveolitis; leukaemia; ICAM-R; VCAM-1; anti-inflammatory;
XX KW intracellular cell adhesion molecule; vascular cell adhesion molecule;
XX KW locomotor recovery; locomotor damage; locomotor impairment;
XX KW autonomic dysfunction; sensory dysfunction; spinal cord injury.
OS Homo sapiens.
XX PN WC200230980-A2.
XX PD 18-APR-2002.
XX PF 15-OCT-2001; 2001WO-US032059.
XX PR 13-OCT-2000; 2000US-00689307.
XX PA (ICOS-) ICOS CORP.
XX PI Gallatin WM, Van Der Vliet M;
XX DR WPI; 2002-463260/49.
XX PT Use of an anti-alpha-d monoclonal antibodies for promoting locomotor
XX PT recovery, inhibiting locomotor damage, limiting locomotor impairment, or
XX PT limiting autonomic and sensory dysfunction following spinal cord injury.
XX PS Example 5; Page 191-194; 270pp; English.
XX CC The invention relates to promoting locomotor recovery, inhibiting
XX CC locomotor damage, limiting locomotor impairment, or limiting autonomic
XX CC and sensory dysfunction following spinal cord injury by administering an
XX CC anti-alpha d (Beta2 integrin alpha2 subunit) monoclonal antibody to a
XX CC spinal cord injury victim. The method also involves the use of a ligand
XX CC selected from ICAM-R or VCAM-1 (intracellular cell adhesion molecule,
XX CC vascular cell adhesion molecule). The method is useful for promoting
XX CC locomotor recovery, inhibiting locomotor damage, limiting locomotor
XX CC impairment, or limiting autonomic and sensory dysfunction following

CC spinal cord injury. In particular, the spinal cord injury comprises
 CC compression of the spinal cord. The antibodies are also useful for
 CC reducing inflammation at the site of a central nervous system injury. The
 CC specification also details the identification of Beta2 integrin alpha
 CC cDNAs and proteins, for use in raising the antibodies. Beta2 integrins
 CC are implicated in diseases such as IAD (leukocyte adhesion deficiency,
 CC inflammatory response, diabetes, multiple sclerosis, arthritis, graft
 CC atherosclerosis, inflammatory bowel disease, Crohn's disease, ulcerative
 CC colitis, immune complex alveolitis and leukaemia. The present sequence is
 CC a Beta2 integrin alpha subunit sequence included for comparison with the
 CC Beta2 integrin alpha protein sequences
 XX
 SQ Sequence 1153 AA;

Query Match 100.0%; Score 5956; DB 5; Length 1153;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRVLLTALFLCHGFNLDENATFQENARGFGQSVVQLGSRVVVVGAPQEIIVANQR 60
 DB 1 MALRVLLTALFLCHGFNLDENATFQENARGFGQSVVQLGSRVVVVGAPQEIIVANQR 60
 QY 61 GSLYOCDSYSGCEPRLOVPEAVNMSGLSLAATTSPQLLAGPTVHOTCSNTYVK 120
 DB 61 GSLYOCDSYSGCEPRLOVPEAVNMSGLSLAATTSPQLLAGPTVHOTCSNTYVK 120
 QY 121 GLCFLFGSNLRQPPKPPALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQL 180
 DB 121 GLCFLFGSNLRQPPKPPALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQL 180
 QY 181 KSKTLFSLMOYSEPRHFTPEKONNPNPSLVKPTTOLLGRTHATGKVKVRELFN 240
 DB 181 KSKTLFSLMOYSEPRHFTPEKONNPNPSLVKPTTOLLGRTHATGKVKVRELFN 240
 QY 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPSEKSRQEL 300
 DB 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPSEKSRQEL 300
 QY 301 NTIASKPPDRHVFQNNFPAKTIQNLREKIFAEGTGTGSSSFEHMSOEGESAALT 360
 DB 301 NTIASKPPDRHVFQNNFPAKTIQNLREKIFAEGTGTGSSSFEHMSOEGESAALT 360
 QY 361 SNGPLLSTVGSYDAGGVFLYTSKESKTFINTRVDSMDNDAYLGVAAILLRNVQSIV 420
 DB 361 SNGPLLSTVGSYDAGGVFLYTSKESKTFINTRVDSMDNDAYLGVAAILLRNVQSIV 420
 QY 421 LGAPRYQHIGLVAMFRQNTGMESNANVKTGTOIGAYFGASLCSVDVDSNGSTDLLVIGAP 480
 DB 421 LGAPRYQHIGLVAMFRQNTGMESNANVKTGTOIGAYFGASLCSVDVDSNGSTDLLVIGAP 480
 QY 481 HYTEQTRGGQSVCPPLRGORARWQCDVLYGEOGPWGRFGAALTVLGDVNGDKLTDVA 540
 DB 481 HYTEQTRGGQSVCPPLRGORARWQCDVLYGEOGPWGRFGAALTVLGDVNGDKLTDVA 540
 QY 541 IGAPGEDNRGAVYLPFGTSGSGISPSHSORVAGSKLSPRLQYFGOSLGGQDLTMDGLV 600
 DB 541 IGAPGEDNRGAVYLPFGTSGSGISPSHSORVAGSKLSPRLQYFGOSLGGQDLTMDGLV 600
 QY 601 DLTVGAQGHVLLRSQPLRVKAIMPNPREVARNVFECDQVVKKEAGEVRVCLHVOK 660
 DB 601 DLTVGAQGHVLLRSQPLRVKAIMPNPREVARNVFECDQVVKKEAGEVRVCLHVOK 660
 QY 661 STYDRLRREGQIQSVVTYDLALDSGRPHSRVAFNETKNSTRRTQVGLTQTCTELKQLP 720
 DB 661 STYDRLRREGQIQSVVTYDLALDSGRPHSRVAFNETKNSTRRTQVGLTQTCTELKQLP 720
 QY 721 NCIEDPVSIVLRNLSVGTPLSAGNLRPVLAEDAORLFTALPPEKNCNDNICODD 780
 DB 721 NCIEDPVSIVLRNLSVGTPLSAGNLRPVLAEDAORLFTALPPEKNCNDNICODD 780
 QY 781 LSITTFPMSLDCLVVGPREFNTVTVRNDGDSYRTQVTFPPDLSTYRKVSTLQNRKS 840
 DB 781 LSITTFPMSLDCLVVGPREFNTVTVRNDGDSYRTQVTFPPDLSTYRKVSTLQNRKS 840

QY 841 QRSWRLACSSASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLIKA 900
 DB 841 QRSWRLACSSASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLIKA 900
 QY 901 NVTSENMPRTNKTETFOLELPVKYAVVMVTSHGVSSTKYLNFPTASENTSRVMQHOYQVSN 960
 DB 901 NVTSENMPRTNKTETFOLELPVKYAVVMVTSHGVSSTKYLNFPTASENTSRVMQHOYQVSN 960
 QY 961 LGQSLDPLISLVFLVPVRLNQTIVWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPV 1020
 DB 961 LGQSLDPLISLVFLVPVRLNQTIVWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPV 1020
 QY 1021 VNCISAVQRIQCDIPFFGIQOEENATLKGNSPDWYIKTSHNHLIVSTAEIILFNDSVF 1080
 DB 1021 VNCISAVQRIQCDIPFFGIQOEENATLKGNSPDWYIKTSHNHLIVSTAEIILFNDSVF 1080
 QY 1081 TLLPGQCAFVRSQTEIKVPEPEVNPPLIVGSSVGGLLILALITAALYKLGFFKROVKD 1140
 DB 1081 TLLPGQCAFVRSQTEIKVPEPEVNPPLIVGSSVGGLLILALITAALYKLGFFKROVKD 1140
 QY 1141 MMSEGGPPGAPQP 1153
 DB 1141 MMSEGGPPGAPQP 1153
 RESULT 5
 AA014428
 ID AA014428 standard; protein; 1153 AA.
 XX AC AA014428;
 XX DT 03-MAY-2002 (first entry)
 DE Integrin Mac-1 alpha subunit.
 KW Mac-1; integrin alpha subunit; variant integrin inserted domain protein;
 KW open conformation; integrin related inflammatory disorder;
 KW integrin related immunological disorder; rheumatoid arthritis; ischaemia;
 KW reperfusion; hypovolemic shock; infarction; cerebral shock;
 KW viral infection; cancer; gene therapy; vaccine;
 KW bioactive agent screening.
 XX Unidentified.
 XX WO200204521-A2.
 XX PD 17-JAN-2002.
 XX PF 09-JUL-2001; 2001WO-US021805.
 XX PR 07-JUL-2000; 2000US-0216600P.
 XX PA (CALY) CALIFORNIA INST OF TECHNOLOGY.
 XX PA (BLOO-) CENT BLOOD RES.
 XX PI Springer T;
 XX WPI; 2002-148167/19.
 XX New integrin I domain protein having alteration in at least 2
 XX noncontiguous regions and exits in an open conformation, useful for
 XX treating, preventing or suppressing inflammatory or immunological
 XX disorders.
 XX Example 1; Fig 1F; 90pp; English.
 XX The invention comprises structurally biased variant integrin inserted (I)
 XX domain proteins, wherein the alterations to the protein occur in at least
 XX two noncontiguous regions. Specifically the variant integrin I domain
 XX proteins are structurally biased to exist in the open conformation,
 XX thereby altering the binding ability of the protein. The invention also
 XX comprises nucleic acids encoding the variant integrin I domain proteins.

CC The integrin I domain proteins and nucleic acids are useful for treating,
CC preventing or suppressing integrin related inflammatory and immunological
CC disorders (e.g. rheumatoid arthritis). The variant integrin I domain
CC proteins and nucleic acids can also be used for treating: ischaemia/
CC reperfusion (e.g. hypovolemic shock); infarction; cerebral ischemia/
CC infection; and cancer. The variant integrin I domain nucleic acids and
CC proteins may be used in gene therapy, as vaccines and to screen for
CC bioactive agents. The present amino acid sequence represents the Mac-1
CC alpha subunit of integrin
XX
SQ Sequence 1153 AA;

Query Match 100.0%; Score 5956; DB 5; Length 1153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRVLLTALTLCCHGNLTENAMTFOENARGGQSWQLOGRVVGAPQIVANOR 60
DB 1 MALRVLLTALTLCCHGNLTENAMTFOENARGGQSWQLOGRVVGAPQIVANOR 60
QY 61 GSLYQCDYSTGSCPEIRLOVPVEAVNMSGLSLAATTSPQLLACGPTVHQTSENTRYK 120
DB 61 GSLYQCDYSTGSCPEIRLOVPVEAVNMSGLSLAATTSPQLLACGPTVHQTSENTRYK 120
QY 121 GLCLFSGNLRQOPKPEALRGCPQEDSDIAFLIDGSGSIIIPHDPRMKEFVSTWQEL 180
DB 121 GLCLFSGNLRQOPKPEALRGCPQEDSDIAFLIDGSGSIIIPHDPRMKEFVSTWQEL 180
QY 181 KSKTFLSLMQYSEEFRIHTTFKFNQNNPRSLVPEITQLLGRTHATGIRKVVRELFN 240
DB 181 KSKTFLSLMQYSEEFRIHTTFKFNQNNPRSLVPEITQLLGRTHATGIRKVVRELFN 240
QY 241 ITNGARKNPKILVITDCKPGDPLGYEDVPEADREGVIRVYIGVGDAPRSEKROEL 300
DB 241 ITNGARKNPKILVITDCKPGDPLGYEDVPEADREGVIRVYIGVGDAPRSEKROEL 300
QY 301 NTIASKPRDHVFOVNNFEALKTIQNLREKIPIAIECTQTGSSSPHEHMSQGFSAIT 360
DB 301 NTIASKPRDHVFOVNNFEALKTIQNLREKIPIAIECTQTGSSSPHEHMSQGFSAIT 360
QY 361 SNGPLSTVSGSYDVGWGVFLYTSKEKSTFTNTRVDSMDNDVLYGAAAILLNRYQSLV 420
DB 361 SNGPLSTVSGSYDVGWGVFLYTSKEKSTFTNTRVDSMDNDVLYGAAAILLNRYQSLV 420
QY 421 LGAPRYOHIGLVAMFRONTGWBSNANVKGTOIGAYFGASLCSDVDNSGSTDVLILGAP 480
DB 421 LGAPRYOHIGLVAMFRONTGWBSNANVKGTOIGAYFGASLCSDVDNSGSTDVLILGAP 480
QY 481 HYYEQTRGGQVSVCLPRGQARWQCDVLYGEGQPGWGFGAALTVLGVDVNGDKLTDVA 540
DB 481 HYYEQTRGGQVSVCLPRGQARWQCDVLYGEGQPGWGFGAALTVLGVDVNGDKLTDVA 540
QY 541 IGAPGEDNRGAVYLFHGTSGSGISPSHSORIGSKLSPLQYFGQSLSGQDLTMDGLV 600
DB 541 IGAPGEDNRGAVYLFHGTSGSGISPSHSORIGSKLSPLQYFGQSLSGQDLTMDGLV 600
QY 601 DLTVGAQGHVLLRSQVPLRVKAIMFNPREVARNVFECDNQVVKGEAGEVSVCLHVQK 660
DB 601 DLTVGAQGHVLLRSQVPLRVKAIMFNPREVARNVFECDNQVVKGEAGEVSVCLHVQK 660
QY 661 STRDLREGQVSVTVLDLSDGRPHSRVFNFTKSTRTQVGLTCTCTKLQLP 720
DB 661 STRDLREGQVSVTVLDLSDGRPHSRVFNFTKSTRTQVGLTCTCTKLQLP 720
QY 721 NCIEDPVSVPLRVNLSVGLTPLSAFGNLRLVLAEDAQRFTALFPPEKNGCDNICQDD 780
DB 721 NCIEDPVSVPLRVNLSVGLTPLSAFGNLRLVLAEDAQRFTALFPPEKNGCDNICQDD 780
QY 781 LSITFSFMSLDCLVVGPREPNVTVVRNDEDSYRTQVTFPPPLDLSYKRVSTLQNR 840
DB 781 LSITFSFMSLDCLVVGPREPNVTVVRNDEDSYRTQVTFPPPLDLSYKRVSTLQNR 840
QY 841 QRSWLACESASSTEVSGALKSTSCSINHPIFPENSEVTFTNITPDVDSKASLGNKLLKA 900

DB 841 QRSWLACESASSTEVSGALKSTSCSINHPIFPENSEVTFTNITPDVDSKASLGNKLLKA 900
QY 901 NVTSENNPRTNKTTFQLELPVKYAVYVWTSKGVSTKYLNFTASENTRVVMQHOYVSN 960
DB 901 NVTSENNPRTNKTTFQLELPVKYAVYVWTSKGVSTKYLNFTASENTRVVMQHOYVSN 960
QY 961 LQORSPLISLVLVPLVRLNQTIVDRPQVTFSENLSSSTCHTKERLPSHSDFLAELRKAPV 1020
DB 961 LQORSPLISLVLVPLVRLNQTIVDRPQVTFSENLSSSTCHTKERLPSHSDFLAELRKAPV 1020
QY 1021 VNCSTAVCQRIQCDIPPFQIQEEFNATLKGNLSPDWYIKTSHNHLIYSTABILFNDVSF 1080
DB 1021 VNCSTAVCQRIQCDIPPFQIQEEFNATLKGNLSPDWYIKTSHNHLIYSTABILFNDVSF 1080
QY 1081 TLLPQOGAFVRSQETKVEPEVPNPLPLIVGSSVGGLLLALITAAALYKLGFFPKQYKD 1140
DB 1081 TLLPQOGAFVRSQETKVEPEVPNPLPLIVGSSVGGLLLALITAAALYKLGFFPKQYKD 1140
QY 1141 NMSEGGPPGABEQ 1153
DB 1141 NMSEGGPPGABEQ 1153

RESULT 6
ADD25615
ID ADD25615 standard; protein; 1153 AA.
XX
AC ADD25615;
XX
DT 15-JAN-2004 (first entry)
XX
DE Binding domain-immunoglobulin fusion protein-associated protein #85.
XX
KW Binding domain; immunoglobulin; fusion protein; cytostatic;
KW antirheumatic; immunosuppressive; antidiabetic; antithyroid;
KW neuroprotective; hinge region; immunoglobulin heavy chain;
KW CH2 constant region; CH3 constant region; IgG1;
KW antibody dependent cell-mediated cytotoxicity; ADCC; complement fixation;
KW malignant condition; B-cell disorder; melanoma; carcinoma; sarcoma;
KW rheumatoid arthritis; myasthenia gravis; Grave's disease;
KW type I diabetes mellitus; multiple sclerosis; autoimmune disease.
XX
OS Unidentified.
XX
PN US2003118592-A1.
XX
PD 26-JUN-2003.
XX
XX 25-JUL-2002; 2002US-00207655.
XX
PF 17-JAN-2001; 2001US-0367358P.
PR 17-JAN-2002; 2002US-00053530.
PR 03-JUN-2002; 2002US-0385691P.
XX
PA (GENE-) GENE-CRAFT INC.
XX
PI Ledbetter JA, Hayden-Ledbetter MS, Thompson PA;
XX
DR WPI; 2003-801317/75.
XX
XX New binding domain-immunoglobulin fusion protein, useful for treating a
PT subject having or suspected of having a malignant condition or a B-cell
PT disorder, e.g. melanoma, Grave's disease or autoimmune disease.
XX
PS Disclosure; SEQ ID NO 176; 157pp; English.
XX
CC The invention relates to a binding domain-immunoglobulin fusion protein
CC comprising a binding domain polypeptide that is fused to an
CC immunoglobulin hinge region polypeptide, an immunoglobulin heavy chain
CC CH2 constant region polypeptide that is fused to the hinge region
CC polypeptide, and an immunoglobulin heavy chain CH3 constant region
CC polypeptide that is fused to the CH2 constant region polypeptide. The

CC hinge region polypeptide comprises: a wild-type human IgG1 immunoglobulin
 CC hinge region polypeptide; a mutated human IgG1 immunoglobulin hinge
 CC region polypeptide, derived from (a) having 3 or more cysteine residues;
 CC where the mutated human IgG1 immunoglobulin hinge region polypeptide
 CC contains 2 cysteine residues, where the first cysteine is not mutated; a
 CC mutated human IgG1 immunoglobulin hinge region polypeptide, derived from
 CC (a) having 3 or more cysteine residues, where the mutated human IgG1
 CC immunoglobulin hinge region polypeptide contains no more than one
 CC cysteine residue; and a mutated human IgG1 immunoglobulin hinge region
 CC polypeptide, derived from (a) having 3 or more cysteine residues; where
 CC the mutated human IgG1 immunoglobulin hinge region polypeptide contains
 CC no cysteine residues. The binding domain-immunoglobulin fusion protein is
 CC capable of at least one immunological activity comprising antibody
 CC dependent cell-mediated cytotoxicity (ADCC) and complement fixation. The
 CC binding domain polypeptide is capable of specifically binding to an
 CC antigen. Also included are an isolated polynucleotide encoding the
 CC binding domain-immunoglobulin fusion protein, a recombinant expression
 CC construct comprising the polynucleotide operably linked to a promoter),
 CC a host cell transformed or transfected with a recombinant expression
 CC construct, producing the binding domain-immunoglobulin fusion protein, a
 CC pharmaceutical composition comprising the binding domain-immunoglobulin
 CC fusion protein or polynucleotide and a carrier, and treating a subject
 CC having or suspected of having a malignant condition or a B-cell disorder.
 CC The binding domain-immunoglobulin fusion protein is useful for treating a
 CC subject having or suspected of having a malignant condition or a B-cell
 CC disorder, e.g. melanoma, carcinoma or sarcoma, rheumatoid arthritis,
 CC myasthenia gravis, Grave's disease, type 1 diabetes mellitus, multiple
 CC sclerosis or autoimmune disease. The present sequence is a binding domain
 CC -immunoglobulin fusion protein-associated protein sequence. Note: The
 CC sequence data for this patent formed part of the printed specification
 CC and is also available in electronic format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html?DocId=20030118592. The authors have not
 CC identified the sequences in the printed specification by their SEQ ID
 CC number therefore none of the sequences can be explicitly identified.

XX SQ Sequence 1153 AA;

Query Match 100.0%; Score 5956; DB 7; Length 1153;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRVLLTALTLCHGFNLDENAMTFQENARGFQGVVQLGSRVVGAPQEIIVANQR 60
 DB 1 MALRVLLTALTLCHGFNLDENAMTFQENARGFQGVVQLGSRVVGAPQEIIVANQR 60
 QY 61 GSLYCDYSTGCEPIRLQVPEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVK 120
 DB 61 GSLYCDYSTGCEPIRLQVPEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVK 120
 QY 121 GLCFIFGSLNRQOPKFPALRGCPQEDSDIAFLIDGSGSIIPHDPRMKCFVSTVMBQL 180
 DB 121 GLCFIFGSLNRQOPKFPALRGCPQEDSDIAFLIDGSGSIIPHDPRMKCFVSTVMBQL 180
 QY 181 KSKTLFSLMOYSEFRHFTFKEFQNNPNPSLVKPIITOLLGRTHATGIRKVVRELFPN 240
 DB 181 KSKTLFSLMOYSEFRHFTFKEFQNNPNPSLVKPIITOLLGRTHATGIRKVVRELFPN 240
 QY 241 ITNGARKNAFKILVITDGEKFGDPLGVEDVIPEADREGVIRYVIGVDGAFRSEKSRQEL 300
 DB 241 ITNGARKNAFKILVITDGEKFGDPLGVEDVIPEADREGVIRYVIGVDGAFRSEKSRQEL 300
 QY 301 NTIASKPRDRHVFQNNFEALNTIQNLREKIPATEGCTGSSSSFEHMSQEGFSAIT 360
 DB 301 NTIASKPRDRHVFQNNFEALNTIQNLREKIPATEGCTGSSSSFEHMSQEGFSAIT 360
 QY 361 SNGPLLSVGSYDMAGGVFLYTSKSKSTFINNTRVDSMDNDAYLQYAAATILNRVQSLV 420
 DB 361 SNGPLLSVGSYDMAGGVFLYTSKSKSTFINNTRVDSMDNDAYLQYAAATILNRVQSLV 420
 QY 421 LGAPRYQHIGLVAMFRONTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAP 480
 DB 421 LGAPRYQHIGLVAMFRONTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAP 480

QY 481 HYIEQTRGGQVSVCPLEPRGQARWQCDAVLYGEGQGPWGRFGAALTVLGVDVNGDKLTDVA 540
 DB 481 HYIEQTRGGQVSVCPLEPRGQARWQCDAVLYGEGQGPWGRFGAALTVLGVDVNGDKLTDVA 540
 QY 541 IGABEDNRGNVLFHGTSGSGISPSHSQRIAGSKLSPLOYEQSLSGGQDLTMDGLV 600
 DB 541 IGABEDNRGNVLFHGTSGSGISPSHSQRIAGSKLSPLOYEQSLSGGQDLTMDGLV 600
 QY 601 DLTVGAQGHVLLLRSQPVLRVKAIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQK 660
 DB 601 DLTVGAQGHVLLLRSQPVLRVKAIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQK 660
 QY 661 STRDLREGQIQSVVYTDALDSGRPHSRAVFNTRKSTRTOVGLTQTCETLKLQLP 720
 DB 661 STRDLREGQIQSVVYTDALDSGRPHSRAVFNTRKSTRTOVGLTQTCETLKLQLP 720
 QY 721 NCIEDPVSPIVRLNFSLVGTPLSAFGLNLPVLAEDAQRFLTALFPPEKNCNDNICODD 780
 DB 721 NCIEDPVSPIVRLNFSLVGTPLSAFGLNLPVLAEDAQRFLTALFPPEKNCNDNICODD 780
 QY 781 LSITFSFMSLDCLVVGGRPFNVTVVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQORS 840
 DB 781 LSITFSFMSLDCLVVGGRPFNVTVVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQORS 840
 QY 841 QRSWRLACESASSSTEVSALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKLLKA 900
 DB 841 QRSWRLACESASSSTEVSALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKLLKA 900
 QY 901 NVTSNNMPRTNKTPEOLELPVKVAVYVNTSHGVSTKYLNFTASENTRVWQHOYQVSN 960
 DB 901 NVTSNNMPRTNKTPEOLELPVKVAVYVNTSHGVSTKYLNFTASENTRVWQHOYQVSN 960
 QY 961 LGQBSLPISLVFLVPLNCTVIMDRPOVTFSENLSCTCHTKERLPSHSDFLAELRKAPV 1020
 DB 961 LGQBSLPISLVFLVPLNCTVIMDRPOVTFSENLSCTCHTKERLPSHSDFLAELRKAPV 1020
 QY 1021 VNCISAVCQRIQCDIPFFGIQEBFNATLKGNLSPDWYIKTSHNELLIVSTAELFNDVSF 1080
 DB 1021 VNCISAVCQRIQCDIPFFGIQEBFNATLKGNLSPDWYIKTSHNELLIVSTAELFNDVSF 1080
 QY 1081 TLLPGQAFVRSQTETKVEFEPVNPPLIVGSSVGLLILALITAAALKLGPFRQYKD 1140
 DB 1081 TLLPGQAFVRSQTETKVEFEPVNPPLIVGSSVGLLILALITAAALKLGPFRQYKD 1140
 QY 1141 MMSEGGPPGAEPO 1153
 DB 1141 MMSEGGPPGAEPO 1153

RESULT 7
 AAR04136
 ID AAR04136 standard; protein; 1153 AA.
 XX AAR04136;
 AC AAR04136;
 XX AC
 XX 25-MAR-2003 (revised)
 DT 07-SEP-1990 (first entry)
 XX Alpha subunit of Mac-1 leukocyte adhesion receptor.
 XX Mac-1 alpha subunit; Mac-1 alpha/beta heterodimer;
 KW non-specific defence system; integrin gene superfamily.
 XX Synthetic.
 XX Key Location/Qualifiers
 FT Region 1..16
 FT /label= signal peptide
 FT Modified-site 86..88
 FT /label= putative N-glycosylation site
 FT Modified-site 240..242
 FT /label= putative N-glycosylation site
 FT Modified-site 391..393

QY	1141	MMSEGGPPGAEPO	1153	
Db	1141	MMSEGGPPGAEPO	1153	
RESULT 8				
AA07120				
ID	AA07120	standard; protein; 1163 AA.		
XX	AA07120;			
RC	AA07120;			
DT	25-MAR-2003	(revised)		
DT	05-FEB-1991	(first entry)		
XX	p150.95	alpha subunit encoded by clone lambdaX47.		
DE	p150.95	leucocyte adhesion receptor alpha-subunit; hairy cell leukaemia;		
KW	rhinovirus.			
XX	Synthetic.			
OS				
PH	Key	Location/Qualifiers		
FT	Region	1..19		
FT	Region	/label= signal peptide		
FT	Region	20..44		
FT	Modified-site	/label= N-terminus		
FT	Modified-site	61..63		
FT	Modified-site	/label= glycosylation site		
FT	Modified-site	89..91		
FT	Modified-site	/label= glycosylation site		
FT	Modified-site	385..387		
FT	Modified-site	/label= glycosylation site		
FT	Modified-site	392..394		
FT	Modified-site	/label= glycosylation site		
FT	Modified-site	697..699		
FT	Modified-site	/label= glycosylation site		
FT	Modified-site	735..737		
FT	Modified-site	/label= glycosylation site		
FT	Modified-site	899..901		
FT	Modified-site	/label= glycosylation site		
FT	Modified-site	904..906		
FT	Modified-site	/label= glycosylation site		
FT	Modified-site	939..941		
FT	Modified-site	/label= glycosylation site		
FT	Modified-site	1050..1052		
FT	Modified-site	/label= glycosylation site		
FT	Modified-site	1108..1133		
FT	Domain	/label= transmembrane		
XX	WO9010646-A.			
XX	20-SEP-1990.			
XX	09-MAR-1990;	90WO-US0001257.		
XX	09-MAR-1990;	90WO-US0001257.		
XX	(DAND)	DANA FARBER CANCER INST INC.		
PA	Corbi AA,	Springer TA;		
PI	WPI;	1990-304985/40.		
DR	N-PSDB;	AAQ06068.		
XX	Treatment of viral esp. rhino-viral infection - by admin. of alpha sub-			
PT	unit of p150.95 cell surface adhesion receptor, opt. together with a beta			
PT	chain of CD-18 family.			
XX	Disclosure; Fig 3; 59pp; English.			
PS	Clone lambda X47 was isolated from a cDNA library constructed from total			
XX	RNA extracted from phorbol myristate acetate stimulated HL-60			
CC	myelomonocytic cells. The library was screened with oligonucleotide			
CC				

CC	probes based on tryptic peptide fragments of p150.95. The sequence can be			
CC	attached to appropriate control elements and expressed in prokaryotic and			
CC	eukaryotic cells. The protein can be used to treat or prevent rhinoviral			
CC	infection because it interacts with ICAM-1 and inhibits cell-virus			
CC	attachment. It can also be used as an anti-inflammatory agent. See also			
CC	AAQ06063-4, AAQ06068, AAR07121-8 and AAR07152-6. (Updated on 25-MAR-2003			
CC	to correct PR field.) (Updated on 25-MAR-2003 to correct PA field.)			
XX	Sequence 1163 AA;			
XX				
QY	Query Match	58.7%; Score 3494; DB 2; Length 1163;		
QY	Best Local Similarity	61.2%; Pred. No. 4e-282;		
QY	Matches	698; Conservative 138; Mismatches 299; Indels 6; Gaps 4;		
QY	5	VLLLTALTLCGFLNLTENAMTFOENARGFGQSVQVQGGVVRVVGAPQEIIVAAQGRSLY 64		
Db	8	LLLTALATSLGFLNLTDELTAFRVDGAGFQSVQVQANSVVRVVGAPQKITAAQNTGGLY 67		
QY	65	QCYSTGSCPIRLQVPVAVNMSLGLSLAATSPQOLLACGPTVHOTCSENTVKGLCF 124		
Db	68	QCYSTGACBPGLQVPVAVNMSLGLSLAATSPQOLLACGPTVHOTCSENTVKGLCF 127		
QY	125	LFGSNLRQQPKPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKFVSVTVMQLKSK 184		
Db	128	LLGPT--QLTORLPVSRQCPQEDSDIAFLIDGSGSIIPHDPRMKFVSVTVMQLKSK 185		
QY	185	TLFSLMQVSEBFRTHPTFKFQNNPNRSLVKPITQLLGRHTHTATGRKVVRELFNTNG 244		
Db	186	TOPSLMQVSEBFRTHPTFKFQNNPNRSLVKPITQLLGRHTHTATGRKVVRELFNTNG 245		
QY	245	ARKNAFLAVITDGEKFGDPLGYEDVPEADREGVIRYVIGVDGAPSEKSRQBLNTIA 304		
Db	246	ARRDATKILIVITDGEKFGDPLGYEDVPEADREGVIRYVIGVDGAPSEKSRQBLNTIA 305		
QY	305	SKPPRDHVQVNNFEALKTIONLREKIFAIEGTQTSSTSSFEHEMSQEGSAITSNP 364		
Db	306	SKPSQEHIFKVEDFDALKDIONLREKIFAIEGTQTSSTSSFEHEMSQEGSAITSNP 365		
QY	365	LLSTVGSYDAGVGLYTSKEKSTFINMTRVDSMDNAYLGYAAAILRNVSQSLVLCAP 424		
Db	366	VLGAVGSFTSGGAPLYPPNNSPTFIMNSQENVMRDSILGYSTELAWKGVQSLVLCAP 425		
QY	425	RYQHIGLVAMFRQNTGMWESNANTKGTQIGAYFGASCLSDVDVDSNGSTDLVLGAPHYYE 484		
Db	426	RYQHTGKAVITQVSRQWEMKAEVTQIGSYFGASCLSDVDVDSNGSTDLVLGAPHYYE 485		
QY	485	QTRGQSVSCPLPRGQARWOCDAVLGEQOQPHRGGAALTVLGDVNGDKLTQVLCAP 544		
Db	486	QTRGQSVSCPLPRGWR--RWCDAVLGEQOQPHRGGAALTVLGDVNGDKLTQVLCAP 544		
QY	545	GEEDNRGAVLPHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTWDGLVDLTV 604		
Db	545	GEENRGAVLPHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTWDGLVDLTV 604		
QY	605	GAQGHVLLRSQVLRVKAIMEFNPREVARNVFECDQVVKKGAGEVVRVCLHVKQSTRD 664		
Db	605	GARGQVLLRTRPVLWGVSMQFIPAEIPRSFAFECREQVSEQTLVQSNICLYIDKRSKN 664		
QY	665	RLREGQIOSVVTYDLALDQSRPHSAVENETKSNTRROTQVLGLTQTCETLKLQPNCE 724		
Db	665	LLGSRDQSSVTLDLALDPGRLSPRATFOETRNLSRVRVLGLKAHCENFLLPSQV 724		
QY	725	DFVSPVILRLNPLSVGTPLSAFQNLRLPVLADQRLFTALPFFKNCQNDNICQDLISIT 784		
Db	725	DSVTPTTLRLNPLSVGTPLSAFQNLRLPVLADQRLFTALPFFKNCQNDNICQDLISIT 784		
QY	785	PSFSLDCLVGGPREFVNTVRNDEDSYRTQVTFPFLDLISYRKVYSTIQNQRSSQ 844		
Db	785	PSFPGKSLVGSNLELNAFVWVWMDGSDSYTITFSPHAGLSRYVAGQKQQLKSL 844		
QY	845	RLACBSASTSVGALKSTSCSINHPIPPENSEVFNITFDVDSKASLGNKLLKANVTS 904		
Db	845	HLTCDSA--PVASQGTWSTSCSINHPIPPENSEVFNITFDVDSKASLGNKLLKANVTS 902		

This sequence represents a human beta-integrin CD11c subunit which is used to describe a method for identifying compounds that modulate the interaction of the beta-integrin alpha-d subunit with a binding partner of alpha-d which involves contacting an alpha-d polypeptide with an alpha-d binding partner, one of which is immobilised and the other of which is labelled, in the presence of a test compound, and determining if the compound affects binding between the alpha-d polypeptide and alpha-d binding partner, where the alpha-d polypeptide is alpha-d or its fragment comprising the cytoplasmic, transmembrane or extracellular domain of alpha-d. Compounds that modulate alpha-d binding could be used to treat diseases such as type-1 diabetes, atherosclerosis, multiple sclerosis, asthma, psoriasis, lung inflammation, acute respiratory distress syndrome and rheumatoid arthritis

Db 963 RDLPSVINFVPELVNQAEVWMDVSEVHPQNPFLSCSEKIAAPPDFLAHIQKNPVLDC 1022

QY 1024 STAVCORIQCDIPPTGIGOEENFATLKGMLSPDWYIKTSHNHLIIIVTARILFNDSVFTLL 1083

Db 1023 STAGCLRFCDVPSVQVEELDTLKGMLSGFWVQIILQKRVSVSVVZAEIIFDTSVYSQL 1082

QY 1084 PQGAFVRSQSTKVEPPEVNPPLPLIVGSSVGGLLILALITAAALYKLGFFKQYQYKMS 1143

Db 1083 PQGEAFMAQTTIVLEKYKVNPIPLIVGSSIGGLLLALITAVLYKVGFFKQYKEME 1142

QY 1144 E 1144

Db 1143 E 1143

RESULT 10

ID AAB07361 standard; protein; 1163 AA.

XX AAB07361;

DT 17-JAN-2001 (first entry)

DE Human CD11c protein sequence.

KW Human; macrophage infiltration inhibition; alpha d integrin; leukocyte integrin; Leu-CAM; leukointegrin; immune response; inflammation; leukocyte adhesion deficiency; IAD; Type I diabetes; atherosclerosis; multiple sclerosis; asthma; psoriasis; beta2 integrin; lung inflammation; acute respiratory distress syndrome; Crohn's disease; rheumatoid arthritis; central nervous system injury; CD11c.

XX Homo sapiens.

XX WO200029446-A1.

XX 25-MAY-2000.

XX 16-NOV-1999; 99WO-US027139.

XX 16-NOV-1998; 98US-00193043.

XX 08-JUL-1999; 99US-00350259.

XX (ICOS-) ICOS CORP.

XX Gallatin MM, Van Der Vieren M;

XX WPI; 2000-397751/33.

PT Use of novel anti-alpha integrin d monoclonal antibodies to inhibit macrophage infiltration and reduce inflammation at central nervous system injury sites.

PS Example 5; Fig 1; 270pp; English.

XX Integrins are a class of membrane-associated molecules that participate in cellular adhesion. Integrins are made up of an alpha subunit and a beta subunit. One class of human integrins are restricted to expression in white blood cells and have a common beta2 subunit: the leukocyte integrins, Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins have an important role in immune and inflammatory responses. The present protein sequence is the human integrin alpha subunit CD11c. This sequence was used in an alignment to identify a novel beta2 integrin alpha subunit: alpha d (AAG6014 and AAB07359). The present sequence has approximately 56% identity to the protein sequence of alpha d. The Alpha d gene and protein may be useful in therapy for diseases linked to aberrant alpha d function e.g. Type I diabetes, atherosclerosis, multiple sclerosis, asthma, psoriasis, lung inflammation, acute respiratory distress syndrome, rheumatoid arthritis and leukocyte adhesion deficiency (IAD). In addition, anti-alpha d monoclonal antibodies may be used in the inhibition of macrophage infiltration at the site of a central nervous system injury. The monoclonal antibodies can also be used to detect and diagnose Crohn's disease

XX SQ Sequence 1163 AA;

Query Match 58.4%; Score 3480; DB 3; Length 1163;
Best local similarity 61.1%; Pred. No. 6e-281;
Matches 697; Conservative 144; Mismatches 294; Indels 6; Gaps 4;

QY 5 VLLTALTCLGCHFLDNTENAMTFOENARGCQSVVLOGSVVVGAEIVAAQNSGLY 64

Db 8 LLLFTALATSLGCHFLDTEBLTAFVDSAGFGDSVVQYANSMVVVGAPQKIIAAHQIGLEY 67

QY 65 QCDYSTGSCBPIRLQVPEAVNMSLGLSLAATTSPQALLACGPTVHTQCSNTYKGLCF 124

Db 68 QCGYSTGACBPIGLQVPEAVNMSLGLSLASTTSPQALLACGPTVHHCGRMYLTGLCP 127

QY 125 LFGNLRQOQKFPPEALRGCPQSDIAFLIDGSGSIIPHDFRMKEFVSTVMQLKSK 184

Db 128 LLGPT--QLTQRLVSRQECPEQDQIVFLIDGSGSSSRNEATMNFVRAVISQFORPS 185

QY 185 TLPISLMQVSEBFTHFTFKEPNNENPRSLVKPITQLLGRTHATGKRVVRELFINITMG 244

Db 186 TQSLMQPSNKFQTHFTFEFRFTSNPLSLASVHQLGQFTYTATATQNVVHRLPHASVG 245

QY 245 ARKNAFKILVITDGEKFGDPLGYEDVIPADREGVIRYVIGVDGAPSEKRSKRELNTIA 304

Db 246 ARDAIKILVITDGEKFGDPLGYEDVIPADREGVIRYVIGVDGAPSEKRSKRELNTIA 305

QY 305 SKPRDRHVFQVNNFEALKTIONLREKIFATEGTCTGSSSPSEHEMSEOGFSAATISNGP 364

Db 306 SKPSQEHIFKVEDFDALKDIONLREKIFATEGTCTGSSSPSEHEMSEOGFSAVTPDGP 365

QY 365 LLSTVGSYDWAAGVFLYTSKEKSTFINNTRVDSQNDAYLGAAIILNENRQSVILVGLAP 424

Db 366 VLGAUGSFTWGGAFLYPPNMSPTFINNSQENVDNRDYSYGLSTELALWKVQSVLVLGAP 425

QY 425 RYOHIGLVAMFRONTGWMESNANVKGTOIGAYFGASLCSVDVDSNGSDTLVLIGAPHYIE 484

Db 426 RYOHIGKAVIPIQVSRQWRMKAEVIGTQISYFGASLCSVDVDSNGSDTLVLIGAPHYIE 485

QY 485 QTRGGQVSVCPPLRGQRARWCDAVLYGCEQGPWGRFGAALTVLGDVNGDKLTQVVAIGAP 544

Db 486 QTRGGQVSVCPPLRGWR--RWWCDAVLYGCEQGPWGRFGAALTVLGDVNGDKLTQVVAIGAP 544

QY 545 GEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGOSLGGQDLTWDGLVLTIV 604

Db 545 GEEENRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGOSLGGQDLTWDGLVLTIV 604

QY 605 GNOGHVLLLRSQPLVRVKAIMENPREVARNVFECNDQWKGKEAGEVVRVCLHVOKSTRD 664

Db 605 GARGQVLLLRTPVLWVGVSQGFPAEIPRSAFECEQVSEQTLVQSNICLYIDKRSKN 664

QY 665 RLREGQIQSVVYDIALDSGRPHSAVFNENKGNSTRQTQVLGTQTCETKLQLPNCIE 724

Db 665 LLGSRDLQSSVTLDALAPGRLSPRAIPQETKNRSLSRVVLGKAHCENFNLLPSCVE 724

QY 725 DPVSPVLRLNFSNGTPLSAPGNLRPVLAEDAQRLLFTALPFFKNCNDNICODDLST 784

Db 725 DSVIPIIURLFTLVKPKFLAFNRPLMALLAQAQRYFTASLPFEKNCQADHICQDNLGIS 784

QY 785 FSPMSLDCLVWGGPREFNVTVTVRNDGSDSYRTQVTFPPFLDLSTYRKVSTLQNSQSRW 844

Db 785 FSPFGLKSLVGSNLELNAEVMWMDGSDSYGTTITFSGHPAGLSYRYVAEGKQQLRSL 844

QY 845 RLACASASTFVSGALKKTSCTSIHPIPEKSEVFNITFVDSKASLGNKLLKANVTS 904

Db 845 HLTC--CSAPVGSQGTWTSTSCRNHLIPGGAQITFLATFDVSPKAVGLDRLLLIANVSS 902

QY 905 ENNMPTNKTEFQLELPVKYAVYVWVTVSHGVSTKYLNFTAS--ENTSRVMOHOYQSVNLGQ 963

Db 903 ENNIPTSKTIFQLELPVKYAVYVWVSHQPTKYLNFTSESEKESHVAMHRYQVNNLQ 962

QY 964 RSLPISLTVPLVPRNLQTVIWDPRPOVTFSENLSTCTHYKERLPSHSDFLAEARKKAVNC 1023

D5 963 RDLFVSNFVWPVNEQAVMVDVEVSHQPNQSLRCSSEKIPASDFLAHQXNPVLD 1022
 QY 1024 SIACVQRTQCDIPFGIGCEERNATLKGNSLDFWYKTSNHHLLIVSTAEILENDVPTLL 1083
 DB 1023 SIACVQRTQCDIPFGIGCEERNATLKGNSLDFWYKTSNHHLLIVSTAEILENDVPTLL 1082
 QY 1084 PQGAFVRSOTBTKEVPEVNPVPLIVGSSVGLLALITLALYKLGFFKQYKDMMS 1143
 DB 1083 PQGAFVRSOTBTKEVPEVNPVPLIVGSSVGLLALITLALYKLGFFKQYKDMMS 1142
 QY 1144 E 1144
 DB 1143 E 1143

RESULT 11
 ABG61470
 ID ABG61470 standard; protein; 1163 AA.
 XX
 AC ABG61470;
 XX
 DT 27-AUG-2002 (first entry)
 XX
 DE Human Beta2 integrin alphacD11c subunit.
 XX
 KW Beta2 integrin; alphaD subunit; CD11c subunit; CD11b subunit; LAD;
 KW leukocyte adhesion deficiency; inflammatory response; diabetes;
 KW multiple sclerosis; arthritis; graft atherosclerosis; neuroprotective;
 KW inflammatory bowel disease; Crohn's disease; ulcerative colitis;
 KW immune complex alveolitis; leukaemia; ICAM-R; VCAM-1; anti-inflammatory;
 KW intracellular cell adhesion molecule; vascular cell adhesion molecule;
 KW locomotor recovery; locomotor damage; locomotor impairment;
 KW autonomic dysfunction; sensory dysfunction; spinal cord injury.
 XX
 OS Homo sapiens.
 XX
 PN WO200230980-A2.
 XX
 PD 18-APR-2002.
 XX
 PF 15-OCT-2001; 2001WO-US032059.
 XX
 PR 13-OCT-2000; 2000US-00688307.
 XX
 PA (ICOS-) ICOS CORP.
 XX
 PI Gallatin WW, Van Der Vieren M;
 XX
 DR WPI; 2002-463260/49.
 XX
 PT Use of an anti-alpha-d monoclonal antibodies for promoting locomotor
 PT recovery, inhibiting locomotor damage, limiting locomotor impairment, or
 PT limiting autonomic and sensory dysfunction following spinal cord injury.
 XX
 PS Example 5; Page 194-198; 270pp; English.

The invention relates to promoting locomotor recovery, inhibiting
 locomotor damage, limiting locomotor impairment, or limiting autonomic
 and sensory dysfunction following spinal cord injury by administering an
 anti-alpha d (Beta2 integrin alpha2 subunit) monoclonal antibody to a
 spinal cord injury victim. The method also involves the use of a ligand
 selected from ICAM-R or VCAM-1 (intracellular cell adhesion molecule,
 vascular cell adhesion molecule). The method is useful for promoting
 locomotor recovery, inhibiting locomotor damage, limiting locomotor
 impairment, or limiting autonomic and sensory dysfunction following
 spinal cord injury. In particular, the spinal cord injury comprises
 compression of the spinal cord. The antibodies are also useful for
 reducing inflammation at the site of a central nervous system injury. The
 specification also details the identification of Beta2 integrin alphaD
 cDNAs and proteins, for use in raising the antibodies. Beta2 integrins
 are implicated in diseases such as LAD (leukocyte adhesion deficiency,
 inflammatory response, diabetes, multiple sclerosis, arthritis, graft
 atherosclerosis, inflammatory bowel disease, Crohn's disease, ulcerative

CC colitis, immune complex alveolitis and leukaemia. The present sequence is
 CC a Beta2 integrin alpha subunit sequence included for comparison with the
 CC Beta2 integrin alphaD subunit sequence
 XX
 SQ Sequence 1163 AA;
 Query Match 58.4%; Score 3480; DB 5; Length 1163;
 Best Local Similarity 61.1%; Pred. No. 66-281;
 Matches 697; Conservative 144; Mismatches 294; Indels 6; Gaps 4;
 QY 5 VLLLTALTECHGFNLDTENAMTFQENARFGVQSVQLOQSRVVGAPQEIIVANQSGSLY 64
 DB 8 LLLFTALATSLGFLNLTDELTAFRVDGSGFGVQVYANSVWVVGAPQKILANQIGGLY 67
 QY 65 QCDYSTGSCBPIRLQVPEAVNMSLGLSLAATTSPPQLLACGPTVHTQTCSENYVKGCLP 124
 DB 68 QCGYSTGACEPIGLQVPEAVNMSLGLSLAATTSPPQLLACGPTVHTQTCSENYVKGCLP 127
 QY 125 LFGSNLRQQQKPFPEALRCGPQEDSDIAFLIDGSGSIIIPDFFRMKEFTVSTVMEQKKSK 184
 DB 128 LIQPT--QLTQLPVSRQCPQEQDIIVFLIDGSGSISRNFAFMNFAVVISQFQPS 185
 QY 185 TLFSLMQYSEBFRIHFTPKFQNNPNPSLVKPIITQLLGRTHATGIRKVRBELFNITNG 244
 DB 186 TQPSLMQPSNKFTQHTPEEFRTSNPLSLASVHLQLOQFTYTATAIQNVVHRLFHASYG 245
 QY 245 ARKNAFKILAVITDCGKPGDPLGYEDVPEADREGVIRVVGDAFRSEKSRQELNTIA 304
 DB 246 ARDAIKLIVITDGGKGGSDYKDVIPWADAGIIRIYALVGLAFQENSWKELNDIA 305
 QY 305 SKPPRDEVPQVNNFEALTKIQNLREKIPAIETGTQTGSSSSFEHMSQEGFGAAITNSGP 364
 DB 306 SKPSESHIPKVEDFDALDKIQNLREKIPAIETGTQTGSSSSFEHMSQEGFGAAITNSGP 365
 QY 365 LLSTGVSVDWAGVFLYTSKEKSTPINTRVDSNDNDAYLGYAAAILLRNQSILVIGAP 424
 DB 366 VLGAVGFTWSGGAFLYPPNMSPTFINMSQENVDNRDSYLGYSTELALWKGVQSLVIGAP 425
 QY 425 RYOHIGLVAMFRONTGAMESNANVKGTOIGAFGASLCSVDVDSNGSTDLVLIGAPHYYE 484
 DB 426 RYOHIGKAVIFIQVSRQWRKAEVIGTOIGSYFGASLCSVDVDSNGSTDLVLIGAPHYYE 485
 QY 485 QTRGQGVSVCPPLPRGQARWQCDALYQCGOPWGFAGALTVLGDVNGDKLTDVAIGAP 544
 DB 486 QTRGQGVSVCPPLPRGWR--RWCDALYQCGOPWGFAGALTVLGDVNGDKLTDVAIGAP 544
 QY 545 GEDNREGAVTLFHTGTSGSGISPSHSORIASKLSPRLQYFGQSLSGQGLTMDGLVLT 604
 DB 545 GEDNREGAVTLFHTGTSGSGISPSHSORIASKLSPRLQYFGQSLSGQGLTMDGLVLT 604
 QY 605 GAQGHVLLRSOPVLKVAIMEPNPREVARNVPECDQVVKGEAGEVRLVHVKSTRD 664
 DB 605 GARGQVLLRLTRPVLWVGVSQMOPFPAIPRSAFECEBQVVSQTLVQSNICIDYIDKRSKN 664
 QY 665 RLREGQIQSVVTVYDLDALDSGRPHSAVFNKSTNRQTVLGLTQTCETLKLQPNCLIE 724
 DB 665 LLGSRDLQSVTLDLALAPGRSLSPRAIFQETNRSLSRVRLGLKAKCFENLLPSCVE 724
 QY 725 DPVSPVTLRLNLSVLTPLSAFNGRLAPVLAEDAQRLFTALFFPEKRCNDNTCODDLSIT 784
 DB 725 DSVIPILRLNLTFLVGLKPLAFNRLPMLAQAQRYFTASLPFEKRCNGADHICODMLGIS 784
 QY 785 FSPMSLDCLVVGSPRENTVTVRNDGDSYRTVTFPPFLDLSYRKVSTLQVRSQSW 844
 DB 785 FSPGGLKSLVGNLNLNLAFAVWVNDGDSYGTITFSPAGLSYRYVAEQKQGLRSL 844
 QY 845 RLACESASTEVSGALKSTSCSINHPIIPENSEVTNITFDVDSKASLGNKLLKANVTS 904
 DB 845 HLTC--CSAPVGSQGTWSTSCRINHPIIPRGAQITFLATFDVSPKAVGLDRLLIANVS 902
 QY 905 ENNMPTNTEQLELPVAVYVTVVTSVGVSTKYNLFTAS--ENTSRVVMQHOYQVSNLQ 963
 DB 903 ENNIPRISKTIITQLELPVAVYVTVVTSVGVSTKYNLFTAS--ENTSRVVMQHOYQVSNLQ 962

[illegible]

Search completed: June 7, 2004, 17:11:31
Job time : 55.1568 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 17:16:51 ; Search time 39,4378 Seconds

(without alignments)
8225.189 Million cell updates/sec

Title: US-09-902-481B-1

Perfect score: 5956

Sequence: 1 MALRVLLLTALTLCHGFNLD.....FKRQYKDMWSEGGPRGAPQ 1153

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1155919 seqs, 28138677 residues

Total number of hits satisfying chosen parameters: 1155919

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*

```
1: /cgn2_6/ptodata1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata1/pubpaa/PCT_NEW PUB.pep.*
3: /cgn2_6/ptodata1/pubpaa/US06_NEW PUB.pep.*
4: /cgn2_6/ptodata1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata1/pubpaa/US07_NEW PUB.pep.*
6: /cgn2_6/ptodata1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata1/pubpaa/US08_NEW PUB.pep.*
8: /cgn2_6/ptodata1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata1/pubpaa/US09_NEW PUB.pep.*
13: /cgn2_6/ptodata1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata1/pubpaa/US10_NEW PUB.pep.*
17: /cgn2_6/ptodata1/pubpaa/US60_NEW PUB.pep.*
18: /cgn2_6/ptodata1/pubpaa/US60_PUBCOMB.pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5956	100.0	1153	9 US-09-350-259-3	Sequence 3, Appli
2	5956	100.0	1153	10 US-09-902-481A-1	Sequence 1, Appli
3	5956	100.0	1153	10 US-09-891-943-3	Sequence 3, Appli
4	5956	100.0	1153	14 US-10-144-259-30	Sequence 30, Appl
5	5956	100.0	1153	14 US-10-207-655-176	Sequence 176, App
6	5940.5	99.7	1152	9 US-09-945-265-4	Sequence 4, Appli
7	5868	98.5	1137	10 US-09-902-481A-6	Sequence 6, Appli
8	5862	98.4	1137	10 US-09-902-481A-5	Sequence 5, Appli
9	5852	98.3	1137	10 US-09-902-481A-4	Sequence 4, Appli
10	5839	98.0	1137	10 US-09-902-481A-3	Sequence 3, Appli
11	3503	58.8	1163	15 US-10-116-275-204	Sequence 204, App
12	3480	58.4	1163	9 US-09-350-259-4	Sequence 4, Appli
13	3480	58.4	1163	10 US-09-891-943-4	Sequence 4, Appli
14	3455	58.0	1161	9 US-09-350-259-2	Sequence 2, Appli
15	3455	58.0	1161	10 US-09-891-943-2	Sequence 2, Appli

```
16 3439.5 57.7 1161 9 US-09-350-259-99
17 3439.5 57.7 1161 10 US-09-891-943-99
18 3439.5 57.7 1161 9 US-09-350-259-55
19 3270.5 54.9 1161 9 US-09-891-943-55
20 3264 54.8 1161 9 US-09-350-259-53
21 3264 54.8 1161 10 US-09-891-943-53
22 3251.5 54.6 1151 9 US-09-350-259-37
23 3251.5 54.6 1151 10 US-09-891-943-37
24 3243 54.4 1155 9 US-09-350-259-46
25 3243 54.4 1155 10 US-09-891-943-46
26 1848 31.0 369 12 US-10-087-192-1212
27 1563 26.2 1170 9 US-09-945-365-2
28 1559 26.2 1170 12 US-10-261-164-1
29 1535.5 25.8 1223 16 US-10-408-765A-295
30 1361 22.9 1086 16 US-10-408-765A-1871
31 1229.5 20.6 494 9 US-09-350-259-103
32 1229.5 20.6 494 10 US-09-891-943-103
33 1189.5 20.0 413 9 US-09-350-259-101
34 1189.5 20.0 413 10 US-09-891-943-101
35 1163 19.5 1179 14 US-10-177-550-2
36 1163 19.5 1179 15 US-10-173-551-2
37 1123.5 18.9 1179 12 US-09-918-715-250
38 1110 18.6 1188 15 US-10-291-265-810
39 1108.5 18.6 1151 10 US-09-984-130-103
40 1108.5 18.6 1151 10 US-09-836-353A-103
41 1106 18.6 1188 15 US-10-291-265-338
42 1102.5 18.5 1189 10 US-09-984-130-35
43 1102.5 18.5 1189 10 US-09-836-353A-35
44 1102.5 18.5 1189 12 US-10-262-839-4
45 1087.5 18.3 1181 12 US-10-211-462-187
```

ALIGNMENTS

RESULT 1

```
US-09-350-259-3
; Sequence 3, Application US/09350259
; Patent No. US20020062008A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20020062008A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350,259
; CURRENT FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: 09/193,043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-259-3
```

Query Match 100.0%; Score 5956; DB 9; Length 1153;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 MALRVLLLTALTLCHGFNLDTENAMTFQENARGQSVVQLQSGRVVVGAPQEIIVAAQR 60

Db 1 MALRVLLLTALTLCHGFNLDTENAMTFQENARGQSVVQLQSGRVVVGAPQEIIVAAQR 60

Oy 61 GSLVQCDSTSCSPIRLOVPEAVNMNLSGLSLAATTPPQLLACGPTVHQTCSNTYVK 120

Db	61	GSLYQCDYSTGCEPIRLQVPEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSNTYVK	120
Qy	121	GLCFLFGSNLRQOPKPEALRGCPQEDSDIAFLIDSGSIIIPDPRMKFVSTVMEQL	180
Db	121	GLCFLFGSNLRQOPKPEALRGCPQEDSDIAFLIDSGSIIIPDPRMKFVSTVMEQL	180
Qy	181	KKSKTLFSLMOYSEEFRIHPTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFN	240
Db	181	KKSKTLFSLMOYSEEFRIHPTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFN	240
Qy	241	ITNGARKNAFKILVITDGEKFGDPLGYEDVIPEADREGVIRYVIGVDAPFRSEKSRQEL	300
Db	241	ITNGARKNAFKILVITDGEKFGDPLGYEDVIPEADREGVIRYVIGVDAPFRSEKSRQEL	300
Qy	301	NTIASKPPRDRHVQVNNPEALKTIQNLREKI FAIEGTQTGSSSSFEHMSQEGFSAAIT	360
Db	301	NTIASKPPRDRHVQVNNPEALKTIQNLREKI FAIEGTQTGSSSSFEHMSQEGFSAAIT	360
Qy	361	SNGPLLSTVGSYDWAAGVFLYTSKSKSTFTINMTRVDSMDNDAYLGAAAIILNRVQSLV	420
Db	361	SNGPLLSTVGSYDWAAGVFLYTSKSKSTFTINMTRVDSMDNDAYLGAAAIILNRVQSLV	420
Qy	421	LGAPRYQHIGLVAMFRONTGWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAP	480
Db	421	LGAPRYQHIGLVAMFRONTGWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAP	480
Qy	481	HYTEQTRGGQVSVCPPLRGQRARWQCDVLYGEGQGPWGRFGAALTVLGDVNGDKLTDA	540
Db	481	HYTEQTRGGQVSVCPPLRGQRARWQCDVLYGEGQGPWGRFGAALTVLGDVNGDKLTDA	540
Qy	541	IGAPGEDNRGAVLPHGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGGQDLTMDGLV	600
Db	541	IGAPGEDNRGAVLPHGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGGQDLTMDGLV	600
Qy	601	DLTVGAQGHVLLRSQVPLRVKAIEMFNPREVARNYFECNDQVVKGEAGEVRVCLHVQK	660
Db	601	DLTVGAQGHVLLRSQVPLRVKAIEMFNPREVARNYFECNDQVVKGEAGEVRVCLHVQK	660
Qy	661	STRDLRREGQIQSVVYTDALDSGRPHSAVFNETNSTRROQVGLTQTCETLKLQLP	720
Db	661	STRDLRREGQIQSVVYTDALDSGRPHSAVFNETNSTRROQVGLTQTCETLKLQLP	720
Qy	721	NCIEDPVPVILRLNPLSVGTPLSAFQNLRLPVLAEQAQLFTALPFPEKNGNDNICQDD	780
Db	721	NCIEDPVPVILRLNPLSVGTPLSAFQNLRLPVLAEQAQLFTALPFPEKNGNDNICQDD	780
Qy	781	LSITFMSLDCLVWGSPREFNVTVTRNDGSDSYRTQVTFPPDLDSYRKVSTLQNRS	840
Db	781	LSITFMSLDCLVWGSPREFNVTVTRNDGSDSYRTQVTFPPDLDSYRKVSTLQNRS	840
Qy	841	QRSWRLACBSASSTEVSGALKSTCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKA	900
Db	841	QRSWRLACBSASSTEVSGALKSTCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKA	900
Qy	901	NVTSENMPRTNKTETFOLELPVKYAVVYVTVSHGVSTKYLNFTASENTSRVMOHQYQVSN	960
Db	901	NVTSENMPRTNKTETFOLELPVKYAVVYVTVSHGVSTKYLNFTASENTSRVMOHQYQVSN	960
Qy	961	LGORSPLISLVLVPLVRLNQTVLWDRPQVTFSENLSSTCHTKERLPSHSDFLAEURKAPV	1020
Db	961	LGORSPLISLVLVPLVRLNQTVLWDRPQVTFSENLSSTCHTKERLPSHSDFLAEURKAPV	1020
Qy	1021	VNCIAVCQRIQCDIPFFGQIEEFNATLKGNSLFDWYIKTSHNHLIIVSTABILFNDVSF	1080
Db	1021	VNCIAVCQRIQCDIPFFGQIEEFNATLKGNSLFDWYIKTSHNHLIIVSTABILFNDVSF	1080
Qy	1081	TLLPGQGAFFRSOTETKVEPFEPVNPDLPLIVGSSVCGLLILLIITAALYKLGFFKQYKD	1140
Db	1081	TLLPGQGAFFRSOTETKVEPFEPVNPDLPLIVGSSVCGLLILLIITAALYKLGFFKQYKD	1140
Qy	1141	MMSEGGPPGABPQ	1153

Db	1141	MMSEGGPPGABPQ	1153
RESULT 2			
US-09-902-481A-1			
; Sequence 1, Application US/09902481A			
; Publication No. US20030054440A1			
; GENERAL INFORMATION:			
; APPLICANT: Springer, Timothy			
; APPLICANT: Shimooka, Motomu			
; APPLICANT: Shifman, Julia			
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY			
; FILE REFERENCE: A-70586-1/RPT/RMS/RMK			
; CURRENT APPLICATION NUMBER: US/09/902,481A			
; CURRENT FILING DATE: 2001-07-09			
; PRIOR APPLICATION NUMBER: US 60/216,600			
; PRIOR FILING DATE: 2000-07-07			
; NUMBER OF SEQ ID NOS: 7			
; SOFTWARE: Patent in version 3.1			
; SEQ ID NO 1			
; LENGTH: 1153			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: mat_peptide			
; LOCATION: (17)..()			
; OTHER INFORMATION:			
US-09-902-481A-1			
Query Match			
Best Local Similarity 100.0%; Score 5956; DB 10; Length 1153;			
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	MALRVLLTALTLCCHGFNLDLTENAMTFQENARGFQSVVQLQGSRVVVVGAPOEIVAAQR	60
Db	1	MALRVLLTALTLCCHGFNLDLTENAMTFQENARGFQSVVQLQGSRVVVVGAPOEIVAAQR	60
Qy	61	GSLYQCDYSTGCEPIRLQVPEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSNTYVK	120
Db	61	GSLYQCDYSTGCEPIRLQVPEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSNTYVK	120
Qy	121	GLCFLFGSNLRQOPKPEALRGCPQEDSDIAFLIDSGSIIIPDPRMKFVSTVMEQL	180
Db	121	GLCFLFGSNLRQOPKPEALRGCPQEDSDIAFLIDSGSIIIPDPRMKFVSTVMEQL	180
Qy	181	KKSKTLFSLMOYSEEFRIHPTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFN	240
Db	181	KKSKTLFSLMOYSEEFRIHPTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFN	240
Qy	241	ITNGARKNAFKILVITDGEKFGDPLGYEDVIPEADREGVIRYVIGVDAPFRSEKSRQEL	300
Db	241	ITNGARKNAFKILVITDGEKFGDPLGYEDVIPEADREGVIRYVIGVDAPFRSEKSRQEL	300
Qy	301	NTIASKPPRDRHVQVNNPEALKTIQNLREKI FAIEGTQTGSSSSFEHMSQEGFSAAIT	360
Db	301	NTIASKPPRDRHVQVNNPEALKTIQNLREKI FAIEGTQTGSSSSFEHMSQEGFSAAIT	360
Qy	361	SNGPLLSTVGSYDWAAGVFLYTSKSKSTFTINMTRVDSMDNDAYLGAAAIILNRVQSLV	420
Db	361	SNGPLLSTVGSYDWAAGVFLYTSKSKSTFTINMTRVDSMDNDAYLGAAAIILNRVQSLV	420
Qy	421	LGAPRYQHIGLVAMFRONTGWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAP	480
Db	421	LGAPRYQHIGLVAMFRONTGWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAP	480
Qy	481	HYTEQTRGGQVSVCPPLRGQRARWQCDVLYGEGQGPWGRFGAALTVLGDVNGDKLTDA	540
Db	481	HYTEQTRGGQVSVCPPLRGQRARWQCDVLYGEGQGPWGRFGAALTVLGDVNGDKLTDA	540
Qy	541	IGAPGEDNRGAVLPHGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGGQDLTMDGLV	600
Db	541	IGAPGEDNRGAVLPHGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGGQDLTMDGLV	600

QY	601	DLTVGAQHVLRLRSQPVLRVKAIMEFNPREFNVPFCNDQVVKGEAGEVRVCLHVOK	660
DB	601	DLTVGAQHVLRLRSQPVLRVKAIMEFNPREFNVPFCNDQVVKGEAGEVRVCLHVOK	660
QY	661	STRDLREGQIQSVVTVYDLALDSGRPHSRVAVNFNETNSTRTQTVLGLTQTCETLKLQLP	720
DB	661	STRDLREGQIQSVVTVYDLALDSGRPHSRVAVNFNETNSTRTQTVLGLTQTCETLKLQLP	720
QY	721	NCIEDPVSPIVLRNFSLVGTPLSAFGLNRPVLAEDAQRFTALFPPEKNCNDNICODD	780
DB	721	NCIEDPVSPIVLRNFSLVGTPLSAFGLNRPVLAEDAQRFTALFPPEKNCNDNICODD	780
QY	781	LSITPFSMSLDCLVVGGPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQORS	840
DB	781	LSITPFSMSLDCLVVGGPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQORS	840
QY	841	QSRWELACESASSTEVSGALKSTSCSINHPIPPENSEVTENITFDVDSKASIGNKLLKA	900
DB	841	QSRWELACESASSTEVSGALKSTSCSINHPIPPENSEVTENITFDVDSKASIGNKLLKA	900
QY	901	NVTSENMPRTNKTEFOLELPVKAYVMVTVSHGVSTKYLNFNTASENTSRVMQHQYQVSN	960
DB	901	NVTSENMPRTNKTEFOLELPVKAYVMVTVSHGVSTKYLNFNTASENTSRVMQHQYQVSN	960
QY	961	LGQSLPISLVFLVPVRLNQTVIWRDPQVTFSENLSTCHTKERLPSHSDFLAELRKAPV	1020
DB	961	LGQSLPISLVFLVPVRLNQTVIWRDPQVTFSENLSTCHTKERLPSHSDFLAELRKAPV	1020
QY	1021	VNCSTAVCQRIQCDIPFPGIOEBFNATLKGMLSPDWYIKTSHNHLIIVSTAEIIFNDSVF	1080
DB	1021	VNCSTAVCQRIQCDIPFPGIOEBFNATLKGMLSPDWYIKTSHNHLIIVSTAEIIFNDSVF	1080
QY	1081	TLLPQOGAFVRSQTEKVEPPEVNPPLIIVGSSVGGILLALITAAALYKLGFFKRYKD	1140
DB	1081	TLLPQOGAFVRSQTEKVEPPEVNPPLIIVGSSVGGILLALITAAALYKLGFFKRYKD	1140
QY	1141	MMSEGGPPGAEPQ 1153	
DB	1141	MMSEGGPPGAEPQ 1153	

RESULT 3
US-09-891-943-3
; Sequence 3, Application US/09891943
; Publication No. US20030077278A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; TITLE OF INVENTION: No. US20030077278A1 Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/891,943
; PRIOR FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-891-943-3
Query Match 100.0%; Score 5956; DB 10; Length 1153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1081 TLLPGQAGAFVRSQTETKVPFVEFNPLPLTVGSSVGGLLALLLITAAALYKLGFFKRYQD 1140
Qy 1141 MMSEGGPPGAEPPQ 1153
Db 1141 MMSEGGPPGAEPPQ 1153
RESULT 4
US-10-144-259-30
; Sequence 30, Application US/10144259
; Publication No. US20030109691A1
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; APPLICANT: Li, Rui
; APPLICANT: Xiong, Jian-Ping
; TITLE OF INVENTION: VARIANT INTEGRIN POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 00786-548001
; CURRENT APPLICATION NUMBER: US/10/144,259
; CURRENT FILING DATE: 2002-09-04
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: US 09/758,493
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-259-30
Query Match 100.0%; Score 5956; DB 14; Length 1153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MALRVLLLTALTLCGHNLDENAMTFQENARGFGQSVVQLQGSRVVVGAPQIIVAAQR 60
Db 1 MALRVLLLTALTLCGHNLDENAMTFQENARGFGQSVVQLQGSRVVVGAPQIIVAAQR 60
Qy 61 GSYLQCDYSGSCPIRLQVPEAVNMSLGLSLAATTSPPLLACGPTVHQTCSNTYVK 120
Db 61 GSYLQCDYSGSCPIRLQVPEAVNMSLGLSLAATTSPPLLACGPTVHQTCSNTYVK 120
Qy 121 GLCLFGSLNRQKQPFPEALRGCPQEDSDIAPLIDGSGSIIPHDPRMKKFVSTVWEQL 180
Db 121 GLCLFGSLNRQKQPFPEALRGCPQEDSDIAPLIDGSGSIIPHDPRMKKFVSTVWEQL 180
Qy 181 KKSRTLPSLMQYSEEPRIHETKEFQNNPNRSLVXPDITOLLGRTHTATGIRKVVRELFN 240
Db 181 KKSRTLPSLMQYSEEPRIHETKEFQNNPNRSLVXPDITOLLGRTHTATGIRKVVRELFN 240
Qy 241 ITNGARKNAFKILVWIITDGEKFGDPLGYEDVIPEADREGVIRYVIGVDAFRSEKSRQEL 300
Db 241 ITNGARKNAFKILVWIITDGEKFGDPLGYEDVIPEADREGVIRYVIGVDAFRSEKSRQEL 300
Qy 301 NTIASKPPRHHVQVNNFALKTIONQREKIFAIETQGTGSSSEHEMSQGFSAAIT 360
Db 301 NTIASKPPRHHVQVNNFALKTIONQREKIFAIETQGTGSSSEHEMSQGFSAAIT 360
Qy 361 SNGPLSTVGSYDWAGGVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAILLRNVQSLV 420
Db 361 SNGPLSTVGSYDWAGGVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAILLRNVQSLV 420
Qy 421 LGAPRYOHIGLVANFRONTGWESNANVKTQIGAFPGASLCSDVDVDSNGSTDLVLIGAP 480
Db 421 LGAPRYOHIGLVANFRONTGWESNANVKTQIGAFPGASLCSDVDVDSNGSTDLVLIGAP 480
Qy 481 HYTEOTGGGVSVCPPLPRGORARWQCDVILYGEQGWGFGAALTIVLGDVNGDKLTDVA 540
Db 481 HYTEOTGGGVSVCPPLPRGORARWQCDVILYGEQGWGFGAALTIVLGDVNGDKLTDVA 540
Qy 541 IGAPGEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLV 600

Db 541 IGAPGEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLV 600
Qy 601 DLTVGAQGHVLLRSQPVLRVKAIIMEFPNPREVARNVFECDNVVKGKEAGEVRVCLHVOK 660
Db 601 DLTVGAQGHVLLRSQPVLRVKAIIMEFPNPREVARNVFECDNVVKGKEAGEVRVCLHVOK 660
Qy 661 STRDLREGQIOSVVTYDLALDSGRPHSRVAFNETKSTREOTVGLGTCTCETLKQLP 720
Db 661 STRDLREGQIOSVVTYDLALDSGRPHSRVAFNETKSTREOTVGLGTCTCETLKQLP 720
Qy 721 NCIEDPVSPVILRLNPLSLVGTPLSAFGLNRLPVLAEDAQRLEFALFPPEKNCNDNICODD 780
Db 721 NCIEDPVSPVILRLNPLSLVGTPLSAFGLNRLPVLAEDAQRLEFALFPPEKNCNDNICODD 780
Qy 781 LSITPSFMSLCLVVGGRPREFNVTVVRNDCGDSYRTQVTFPPFLDLSYRVKSTLONORS 840
Db 781 LSITPSFMSLCLVVGGRPREFNVTVVRNDCGDSYRTQVTFPPFLDLSYRVKSTLONORS 840
Qy 841 QRSWLACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKA 900
Db 841 QRSWLACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKA 900
Qy 901 NVTSENMPRNKTEFQLELPVKYAVVMVTSHGVSITKYLNFTASENTRVMOHQYQVSN 960
Db 901 NVTSENMPRNKTEFQLELPVKYAVVMVTSHGVSITKYLNFTASENTRVMOHQYQVSN 960
Qy 961 LGORSPLISLVLVPLVRLNQTIVWDRPOVTFSENLSSTCTKTERLPSHSDFLAELRKAPV 1020
Db 961 LGORSPLISLVLVPLVRLNQTIVWDRPOVTFSENLSSTCTKTERLPSHSDFLAELRKAPV 1020
Qy 1021 VNCSTIACVQRIQCDIPFPGIOEEFNATLKNLSFDWYIKTSHNHLIIIVSTAILFNDVSV 1080
Db 1021 VNCSTIACVQRIQCDIPFPGIOEEFNATLKNLSFDWYIKTSHNHLIIIVSTAILFNDVSV 1080
Qy 1081 TLLPGQAGFVRSQETKVPFVEFNPLPLTVGSSVGGLLALLLITAAALYKLGFFKRYQD 1140
Db 1081 TLLPGQAGFVRSQETKVPFVEFNPLPLTVGSSVGGLLALLLITAAALYKLGFFKRYQD 1140
Qy 1141 MMSEGGPPGAEPPQ 1153
Db 1141 MMSEGGPPGAEPPQ 1153
RESULT 5
US-10-207-655-176
; Sequence 176, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Hayden-Ledbetter, Martha S.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069.401C1
; CURRENT APPLICATION NUMBER: US/10/207,655
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 176
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-207-655-176
Query Match 100.0%; Score 5956; DB 14; Length 1153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MALRVLLLTALTLCGHNLDENAMTFQENARGFGQSVVQLQGSRVVVGAPQIIVAAQR 60
Db 1 MALRVLLLTALTLCGHNLDENAMTFQENARGFGQSVVQLQGSRVVVGAPQIIVAAQR 60
Qy 61 GSYLQCDYSGSCPIRLQVPEAVNMSLGLSLAATTSPPLLACGPTVHQTCSNTYVK 120

```
Db 61 GSIYQCDYSTGSCPIRLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVK 120
Qy 121 GLCFPLGSLNRQOPKPEALRGCGQEDSDIAFLIDGSGSIIPHDFRMRKEFVSTVMEQL 180
Db 121 GLCFPLGSLNRQOPKPEALRGCGQEDSDIAFLIDGSGSIIPHDFRMRKEFVSTVMEQL 180
Qy 181 KSKKTLFSLMQYSEEFRIHFTFKFQNNPNPSLVKPTITQLGRTHATGIRKVVRELFN 240
Db 181 KSKKTLFSLMQYSEEFRIHFTFKFQNNPNPSLVKPTITQLGRTHATGIRKVVRELFN 240
Qy 241 ITNGARKNAFKILVITDGEKGDPLGYEDVTPADREGVIRYVIGVDAFRSEKSRQEL 300
Db 241 ITNGARKNAFKILVITDGEKGDPLGYEDVTPADREGVIRYVIGVDAFRSEKSRQEL 300
Qy 301 NTIASKPPRDHVFQVNNFEALKTIONQIREKIFAIEGTQTGSSSFHEHMSQEGFSAIT 360
Db 301 NTIASKPPRDHVFQVNNFEALKTIONQIREKIFAIEGTQTGSSSFHEHMSQEGFSAIT 360
Qy 361 SNGPLLLSTVGSYDMAGGVFLYTSKESKTFINMTRVDSMDNDAYLGYAAAILRNVRQSLV 420
Db 361 SNGPLLLSTVGSYDMAGGVFLYTSKESKTFINMTRVDSMDNDAYLGYAAAILRNVRQSLV 420
Qy 421 LGAPRYQHILGVAMPFRONTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAP 480
Db 421 LGAPRYQHILGVAMPFRONTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAP 480
Qy 481 HYTEQTRGQGVSVCPPLRGORARWQCDVLYGEGQOPWGRFGAALTVLGDVNGDKLTDVA 540
Db 481 HYTEQTRGQGVSVCPPLRGORARWQCDVLYGEGQOPWGRFGAALTVLGDVNGDKLTDVA 540
Qy 541 IGAPGEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLV 600
Db 541 IGAPGEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLV 600
Qy 601 DLTVGAQGHVLLRSQPVLRVKAIMEFNPREVARNVFECDQVWVKEAGEVRVCLHVQK 660
Db 601 DLTVGAQGHVLLRSQPVLRVKAIMEFNPREVARNVFECDQVWVKEAGEVRVCLHVQK 660
Qy 721 NCTEDVPSPVLRLNLSVGTPLSAFNLRLPVLAEDAQRLFTALPFKXNCNDNICQDD 780
Db 721 NCTEDVPSPVLRLNLSVGTPLSAFNLRLPVLAEDAQRLFTALPFKXNCNDNICQDD 780
Qy 781 LSITPFSMLDCLVWGGPREFNVTVTRNDGDSYRTQVTFPPLDLSTYKYSTLQNRS 840
Db 781 LSITPFSMLDCLVWGGPREFNVTVTRNDGDSYRTQVTFPPLDLSTYKYSTLQNRS 840
Qy 841 QRSWRLACSSASTESVSGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKA 900
Db 841 QRSWRLACSSASTESVSGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKA 900
Qy 901 NVTSENMPRTNKTETFEQLELPVKYAVVMVTVSGVSTKYLNFTASNTSRVMOHQVQSN 960
Db 901 NVTSENMPRTNKTETFEQLELPVKYAVVMVTVSGVSTKYLNFTASNTSRVMOHQVQSN 960
Qy 961 LGQRSPLISLVFLVRLNQTWMDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPV 1020
Db 961 LGQRSPLISLVFLVRLNQTWMDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPV 1020
Qy 1021 VNCISIAVCQRIQCDIPFGIQQEENFATLKGNTLSDFWYIKTSHNLLIVSTABILFNDSP 1080
Db 1021 VNCISIAVCQRIQCDIPFGIQQEENFATLKGNTLSDFWYIKTSHNLLIVSTABILFNDSP 1080
Qy 1081 TLLPGQAFVRSQETKVEPFPVNPFLPLIVGSSVGGHLLALLITAAALKYKGFKKQYKD 1140
Db 1081 TLLPGQAFVRSQETKVEPFPVNPFLPLIVGSSVGGHLLALLITAAALKYKGFKKQYKD 1140
Qy 1141 MMSGEGPPGAEPO 1153
Db 1141 MMSGEGPPGAEPO 1153
```

```
RESULT 5
US-09-945-265-4
; Sequence 4, Application US/09945265
; Patent No. US20020123614A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy A.
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Lu, Chafen
; TITLE OF INVENTION: MODIFIED POLYPEPTIDES STABILIZED IN A
; FILE REFERENCE: CN-002CP
; CURRENT APPLICATION NUMBER: US/09/945,265
; CURRENT FILING DATE: 2001-08-31
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1152
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-945-265-4
```

```
Query Match 99.7%; Score 5940.5; DB 9; Length 1152;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1152; Conservative 0; Mismatches 1; Gaps 1;
```

```
Qy 1 MALRVLLLTALTTLCHGFNLDNTENAMTFQENARGFQSVVQLQSGRVVVGAPQEIIVAANOR 60
Db 1 MALRVLLLTALTTLCHGFNLDNTENAMTFQENARGFQSVVQLQSGRVVVGAPQEIIVAANOR 60
Qy 61 GSIYQCDYSTGSCPIRLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVK 120
Db 61 GSIYQCDYSTGSCPIRLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVK 120
Qy 121 GLCFPLGSLNRQOPKPEALRGCGQEDSDIAFLIDGSGSIIPHDFRMRKEFVSTVMEQL 180
Db 121 GLCFPLGSLNRQOPKPEALRGCGQEDSDIAFLIDGSGSIIPHDFRMRKEFVSTVMEQL 180
Qy 181 KSKKTLFSLMQYSEEFRIHFTFKFQNNPNPSLVKPTITQLGRTHATGIRKVVRELFN 240
Db 181 KSKKTLFSLMQYSEEFRIHFTFKFQNNPNPSLVKPTITQLGRTHATGIRKVVRELFN 240
Qy 241 ITNGARKNAFKILVITDGEKGDPLGYEDVTPADREGVIRYVIGVDAFRSEKSRQEL 300
Db 241 ITNGARKNAFKILVITDGEKGDPLGYEDVTPADREGVIRYVIGVDAFRSEKSRQEL 300
Qy 301 NTIASKPPRDHVFQVNNFEALKTIONQIREKIFAIEGTQTGSSSFHEHMSQEGFSAIT 360
Db 301 NTIASKPPRDHVFQVNNFEALKTIONQIREKIFAIEGTQTGSSSFHEHMSQEGFSAIT 360
Qy 361 SNGPLLLSTVGSYDMAGGVFLYTSKESKTFINMTRVDSMDNDAYLGYAAAILRNVRQSLV 420
Db 361 SNGPLLLSTVGSYDMAGGVFLYTSKESKTFINMTRVDSMDNDAYLGYAAAILRNVRQSLV 420
Qy 421 LGAPRYQHILGVAMPFRONTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAP 480
Db 421 LGAPRYQHILGVAMPFRONTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAP 480
Qy 481 HYTEQTRGQGVSVCPPLRGORARWQCDVLYGEGQOPWGRFGAALTVLGDVNGDKLTDVA 540
Db 481 HYTEQTRGQGVSVCPPLRGORARWQCDVLYGEGQOPWGRFGAALTVLGDVNGDKLTDVA 540
Qy 541 IGAPGEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLV 600
Db 541 IGAPGEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLV 600
Qy 601 DLTVGAQGHVLLRSQPVLRVKAIMEFNPREVARNVFECDQVWVKEAGEVRVCLHVQK 660
Db 601 DLTVGAQGHVLLRSQPVLRVKAIMEFNPREVARNVFECDQVWVKEAGEVRVCLHVQK 660
```

QY 661 STDRLREGIOQSVVTVYDLALDSGRPHSRVAFNETKNSRRTOVQLGLTQTCETLKLQLP 720
DB 660 STDRLREGIOQSVVTVYDLALDSGRPHSRVAFNETKNSRRTOVQLGLTQTCETLKLQLP 719
QY 721 NCIEDPVSPIVLRNLSVGLTSLAFNGLRPVLAEDAQRLFTALPFPFKXCGNDNICQDD 780
DB 720 NCIEDPVSPIVLRNLSVGLTSLAFNGLRPVLAEDAQRLFTALPFPFKXCGNDNICQDD 779
QY 781 LSITFSMSLDCLVVGCPREFNVTVVRNDGDSYRTQVTFPPDLISYRKVSTLQONS 840
DB 780 LSITFSMSLDCLVVGCPREFNVTVVRNDGDSYRTQVTFPPDLISYRKVSTLQONS 839
QY 841 QRSWRLACESASSTEVSGALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGKLLKKA 900
DB 840 QRSWRLACESASSTEVSGALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGKLLKKA 899
QY 901 NVTSENMPRTNKTETQLELPVKYAVVMVTVSHGVSTKYLNFNTASENTSRVMOHOYQVN 960
DB 900 NVTSENMPRTNKTETQLELPVKYAVVMVTVSHGVSTKYLNFNTASENTSRVMOHOYQVN 959
QY 961 LGORSLSPIVLRNLSVGLTSLAFNGLRPVLAEDAQRLFTALPFPFKXCGNDNICQDD 1020
DB 960 LGORSLSPIVLRNLSVGLTSLAFNGLRPVLAEDAQRLFTALPFPFKXCGNDNICQDD 1019
QY 1021 VNCISIAVCORIQDIPFPFGIOBENATLKNLSFDWYIKTSHNHLIVSTABILFNDSPV 1080
DB 1020 VNCISIAVCORIQDIPFPFGIOBENATLKNLSFDWYIKTSHNHLIVSTABILFNDSPV 1079
QY 1081 TLLPGQAFVRSOTETKVEPFPVNPPLPLIVGSSVGGLLMLLITAAALYKLGFFKQYKD 1140
DB 1080 TLLPGQAFVRSOTETKVEPFPVNPPLPLIVGSSVGGLLMLLITAAALYKLGFFKQYKD 1139
QY 1141 MMSEGGPPGAEPQ 1153
DB 1140 MMSEGGPPGAEPQ 1152

RESULT 7

US-09-902-481a-6
; Sequence 6, Application US/09902481a
; Publication No. US2003005440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimada, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902.481a
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent version 3.1
; SEQ ID NO 6
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481a-6

Query Match 98.5%; Score 5868; DB 10; Length 1137;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 17 FNLTENAMTQENARGQSVVOLQSRVTVVGAPOEIVANORGSVYQCDYSTGSCPEI 76
DB 1 FNLTENAMTQENARGQSVVOLQSRVTVVGAPOEIVANORGSVYQCDYSTGSCPEI 60
QY 77 RLOPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGCLCFGLGSLNRQQFQK 136
DB 61 RLOPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGCLCFGLGSLNRQQFQK 120

RESULT 8
US-09-902-481a-5
; Sequence 5, Application US/09902481a

QY 137 PPEALRGCPQEDSDIAFLIDGSGSIIPEHPRMKEFVSTVMEQLKSKTILFSLMOYSBEF 196
DB 136 PPEALRGCPQEDSDIAFLIDGSGSIIPEHPRMKEFVSTVMEQLKSKTILFSLMOYSBEF 180
QY 197 RIHPTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELPNITNGARKNAPKILVVI 256
DB 196 RIHPTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELPNITNGARKNAPKILVVI 240
QY 257 TDGKFGDPLGVEDVPEADREGVIRYVIGVDAPFRSBSKQELNTIASKPRDHFVQVN 316
DB 256 TDGKFGDPLGVEDVPEADREGVIRYVIGVDAPFRSBSKQELNTIASKPRDHFVQVN 300
QY 317 NFEALKTIQONLREKIPAIETGTOTGSSSSFEHKSQEGFSAATTSNGPILLSVGSYDAG 376
DB 316 NFEALKTIQONLREKIPAIETGTOTGSSSSFEHKSQEGFSAATTSNGPILLSVGSYDAG 360
QY 377 GVFLYTSKEKSTFIMTRVDSMDNDAYLGAAAAIILNRVOSLVLGAPRYOHIGLVAMFR 436
DB 376 GVFLYTSKEKSTFIMTRVDSMDNDAYLGAAAAIILNRVOSLVLGAPRYOHIGLVAMFR 420
QY 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLGAPHYBQTRGGQSVSCPL 496
DB 436 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLGAPHYBQTRGGQSVSCPL 480
QY 497 PRGQBARWQCDVAVLYGEGQPGWRFGAALTVDGVNEXLTDVAICAPGEEDNRGAVYLF 556
DB 496 PRGQBARWQCDVAVLYGEGQPGWRFGAALTVDGVNEXLTDVAICAPGEEDNRGAVYLF 540
QY 557 HGTSGSGISPHSOSRIAGSKLSPLRYQFGQSLSGQDQDLMGDLVLTGVAQGHVLLARSQ 616
DB 556 HGTSGSGISPHSOSRIAGSKLSPLRYQFGQSLSGQDQDLMGDLVLTGVAQGHVLLARSQ 600
QY 617 PVLVRKAIEMFNPREVARNVPECNDQVYKGEAGEVRVCLHVOKSTRDLREGQIQSVVT 676
DB 616 PVLVRKAIEMFNPREVARNVPECNDQVYKGEAGEVRVCLHVOKSTRDLREGQIQSVVT 660
QY 677 YDLALDSGRPHSRVAFNETKNSRRTOVQLGLTQTCETLKLQLPNCIEDPVPSPVILRLNF 736
DB 676 YDLALDSGRPHSRVAFNETKNSRRTOVQLGLTQTCETLKLQLPNCIEDPVPSPVILRLNF 720
QY 737 SLVGTPLSAGFNLRLPVLAEDAQRLFTALPFPFKXCGNDNICQDDLSITFSPMSLDCLVVG 796
DB 736 SLVGTPLSAGFNLRLPVLAEDAQRLFTALPFPFKXCGNDNICQDDLSITFSPMSLDCLVVG 780
QY 797 GREFNVTTVVRNDGDSYRTQVTFPPDLISYRKVSTLQONSORSMRLACESASSTEV 856
DB 796 GREFNVTTVVRNDGDSYRTQVTFPPDLISYRKVSTLQONSORSMRLACESASSTEV 840
QY 857 SGALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGKLLKXANVTSENNPRTNKTBF 916
DB 856 SGALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGKLLKXANVTSENNPRTNKTBF 900
QY 917 QLELPVKYAVVMVTVSHGVSTKYLNFNTASENTSRVMOHOYQVSNLQORSLSPLVPLVPV 976
DB 916 QLELPVKYAVVMVTVSHGVSTKYLNFNTASENTSRVMOHOYQVSNLQORSLSPLVPLVPV 960
QY 977 RLNQTVIWDPRQVTFSENLSSSTCHTKERLPSHSDFLAELRKAPVVCNCSIAVCORIQCDIP 1036
DB 976 RLNQTVIWDPRQVTFSENLSSSTCHTKERLPSHSDFLAELRKAPVVCNCSIAVCORIQCDIP 1020
QY 1037 FPGIOBENATLKNLSFDWYIKTSHNHLIVSTABILFNDSPVTLPGQAFVRSQET 1096
DB 1036 FPGIOBENATLKNLSFDWYIKTSHNHLIVSTABILFNDSPVTLPGQAFVRSQET 1080
QY 1097 KVEPFPVNPPLPLIVGSSVGGLLMLLITAAALYKLGFFKQYKMMSEGGPPGAEPQ 1153
DB 1096 KVEPFPVNPPLPLIVGSSVGGLLMLLITAAALYKLGFFKQYKMMSEGGPPGAEPQ 1137

Publication No. US20030054440A1
GENERAL INFORMATION:
APPLICANT: Springer, Timothy
APPLICANT: Shimaoka, Motomu
APPLICANT: Shifman, Julia
APPLICANT: Mayo, Stephen
TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
FILE REFERENCE: A-70586-1/RPT/RMS/RMK
CURRENT APPLICATION NUMBER: US/09/902,481A
CURRENT FILING DATE: 2001-07-09
PRIOR APPLICATION NUMBER: US 60/216,600
PRIOR FILING DATE: 2000-07-07
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn version 3.1
SEQ ID NO 5
LENGTH: 1137
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: synthetic
US-09-902-481A-5

Query Match 98.4%; Score 5862; DB 10; Length 1137;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 17 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVGAPQEIIVANQSGSLYQCDYSTGSCPEI 76
DB 1 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVGAPQEIIVANQSGSLYQCDYSTGSCPEI 60

QY 77 RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHOTSSENTYVKGCLFGLFGLNLRQOQPK 136
DB 61 RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHOTSSENTYVKGCLFGLFGLNLRQOQPK 120

QY 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKPFVSTWMEQLKSKTFLSLMOYSEEF 196
DB 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKPFVSTWMEQLKSKTFLSLMOYSEEF 180

QY 197 RIHFTFKFQNNPNSRSLVKTITQLGRTHATGIRKVVRELFINTEGARKNAFKILVVI 256
DB 181 RIHFTFKFQNNPNSRSLVKTITQLGRTHATGIRKVVRELFINTEGARKNAFKILVVI 240

QY 257 TDGKFGDPLGYEDVPEADREGVIRYVIGVDGDAFRSEKSRQELNTIASKPPRDHVFQW 316
DB 241 TDGKFGDPLGYEDVPEADREGVIRYVIGVDGDAFRSEKSRQELNTIASKPPRDHVFQW 300

QY 317 NFEALKTIQNLREKIFAIEGHTQSSSFEHMSQEGPSAAITNGPLLSITVGSYDRAG 376
DB 301 NFEALKTIQNLREKIFAIEGHTQSSSFEHMSQEGPSAAITNGPLLSITVGSYDRAG 360

QY 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAILRNVRQSLVGLGAPRYQHIGLVAMFR 436
DB 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAILRNVRQSLVGLGAPRYQHIGLVAMFR 420

QY 437 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYETRGQSVVCP 496
DB 421 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYETRGQSVVCP 480

QY 497 PRGQARWOCDAVLVGEQGPWGRFGAALTVDVNGDKLTDVAICAPGEEDNRGAVYLF 556
DB 481 PRGQARWOCDAVLVGEQGPWGRFGAALTVDVNGDKLTDVAICAPGEEDNRGAVYLF 540

QY 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVLTGGAQGHVILLRSQ 616
DB 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVLTGGAQGHVILLRSQ 600

QY 617 PVLRVKALINEFNPREVARNVFCNDQVVKGEAGEVRVCLHWOKSTRDLREGQIQSVVT 676
DB 601 PVLRVKALINEFNPREVARNVFCNDQVVKGEAGEVRVCLHWOKSTRDLREGQIQSVVT 660

QY 677 YDLALDSGRPHSAVENETKSTRRTQVLGLTQTCSTLKLQLPNCIEDPVSPIVLRNLF 736
DB 661 YDLALDSGRPHSAVENETKSTRRTQVLGLTQTCSTLKLQLPNCIEDPVSPIVLRNLF 720

QY 737 SLVGTPLSAFGNLRPVLAEDAQRLFTALPFPEKNCNDNITCQDDLSITTFMSLDCLVWG 796
DB 721 SLVGTPLSAFGNLRPVLAEDAQRLFTALPFPEKNCNDNITCQDDLSITTFMSLDCLVWG 780

QY 797 GPREFNVTVVRNDGDSYETQVTFPEPDLVSRKUSTLQNSORSQSRWLACESASSTEV 856
DB 781 GPREFNVTVVRNDGDSYETQVTFPEPDLVSRKUSTLQNSORSQSRWLACESASSTEV 840

QY 857 SGALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGKLLKXANTSENMPRTNKTET 916
DB 841 SGALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGKLLKXANTSENMPRTNKTET 900

QY 917 QLELPVKIAYVMVTVSHGVSTKYLNFITASENTSRVMOHOYOVSNLQORSIPISLVPLVP 976
DB 901 QLELPVKIAYVMVTVSHGVSTKYLNFITASENTSRVMOHOYOVSNLQORSIPISLVPLVP 960

QY 977 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSSHDFLAELRKA PVVNCIAVCQRIQCDIP 1036
DB 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSSHDFLAELRKA PVVNCIAVCQRIQCDIP 1020

QY 1037 PFGIOEEFNATLKGNSLFDWYIKTSHNHLIVSTABEILFNDVSFTLLPGQAFVRSQTE 1096
DB 1021 PFGIOEEFNATLKGNSLFDWYIKTSHNHLIVSTABEILFNDVSFTLLPGQAFVRSQTE 1080

QY 1097 KVEPPEVPNPLPIVGSVGGLLLLALITAAALYKLGFPKQYKDMMSGGGPPGAE 1153
DB 1081 KVEPPEVPNPLPIVGSVGGLLLLALITAAALYKLGFPKQYKDMMSGGGPPGAE 1137

RESULT 9
US-09-902-481A-4
Sequence 4: Application US/09902481A
Publication No. US20030054440A1
GENERAL INFORMATION:
APPLICANT: Springer, Timothy
APPLICANT: Shimaoka, Motomu
APPLICANT: Shifman, Julia
APPLICANT: Mayo, Stephen
TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
FILE REFERENCE: A-70586-1/RPT/RMS/RMK
CURRENT APPLICATION NUMBER: US/09/902,481A
CURRENT FILING DATE: 2001-07-09
PRIOR APPLICATION NUMBER: US 60/216,600
PRIOR FILING DATE: 2000-07-07
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 1137
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: synthetic
US-09-902-481A-4

Query Match 98.3%; Score 5852; DB 10; Length 1137;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 17 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVGAPQEIIVANQSGSLYQCDYSTGSCPEI 76
DB 1 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVGAPQEIIVANQSGSLYQCDYSTGSCPEI 60

QY 77 RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHOTSSENTYVKGCLFGLFGLNLRQOQPK 136
DB 61 RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHOTSSENTYVKGCLFGLFGLNLRQOQPK 120

QY 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKPFVSTWMEQLKSKTFLSLMOYSEEF 196
DB 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKPFVSTWMEQLKSKTFLSLMOYSEEF 180

QY 197 RIHFTFKFQNNPNSRSLVKTITQLGRTHATGIRKVVRELFINTEGARKNAFKILVVI 256
DB 181 RIHFTFKFQNNPNSRSLVKTITQLGRTHATGIRKVVRELFINTEGARKNAFKILVVI 240

181 RIHFTFEQWNNPRSLIKPIITQLLGRTHATGLRVVRELFTNITNGARKNAFKILFLL 240
257 TDGEKFGDLYEDVPEADREGVIRVVGDAFSEKSEKROELNTTASPRDHVFOVN 316
241 TDGEKFGDLYEDVPEADREGVIRVVGDAFSEKSEKROELNTTASPRDHVFOVN 300
317 NFELAKTIQNLREKIFAIETGOTGSSSSFEHEMSQEGFSAALITNGPLLSVGSYDWAG 376
301 NFELAKTIQNLREKIFAIETGOTGSSSSFEHEMSQEGFSAALITNGPLLSVGSYDWAG 360
377 GVFLYTSKESKSTFNNTRVDSMDNDAYLGAALILNRVQSLVGLGAPRYCHIGLVAMFR 436
361 GVFLYTSKESKSTFNNTRVDSMDNDAYLGAALILNRVQSLVGLGAPRYCHIGLVAMFR 420
437 QNTGMWESNANVKGTQICAYFGASCLSDVDVDSNGSTDLVLIGAPHYETRGQVSCPL 496
421 QNTGMWESNANVKGTQICAYFGASCLSDVDVDSNGSTDLVLIGAPHYETRGQVSCPL 480
497 PRGQARWQCDVLYGEGQPGWGFAGALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556
481 PRGQARWQCDVLYGEGQPGWGFAGALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
557 HGTSGSGISPSHSORIASKLSPLQYFGOSLSGGQDLTMDGLVDTLVGAQGHVLLRSQ 616
541 HGTSGSGISPSHSORIASKLSPLQYFGOSLSGGQDLTMDGLVDTLVGAQGHVLLRSQ 600
617 PVLRVKAIMFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLRBEQIQSVVT 676
601 PVLRVKAIMFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLRBEQIQSVVT 660
677 YDLALDSGRPHSRAVFNETKNSRQTVGLTQTCETLKLQPNCEIDPVSPIVLRNLF 736
661 YDLALDSGRPHSRAVFNETKNSRQTVGLTQTCETLKLQPNCEIDPVSPIVLRNLF 720
737 SLVGTPLSAFNGLRPVLAEDAQRLLFTALFPPEKNCNDNITCQDDLSTTFPSMGLDCLVWG 796
721 SLVGTPLSAFNGLRPVLAEDAQRLLFTALFPPEKNCNDNITCQDDLSTTFPSMGLDCLVWG 780
797 GPREFNVTVVRNDGEDSYRTQVTFPEPLDLSYKVSFLQNSQSRWRLACSSASTEV 856
781 GPREFNVTVVRNDGEDSYRTQVTFPEPLDLSYKVSFLQNSQSRWRLACSSASTEV 840
857 SGALKSTSCSINHPIFPENSEVNTITFDVDSKASLGNKLLKANVTSENMPRTNKTET 916
841 SGALKSTSCSINHPIFPENSEVNTITFDVDSKASLGNKLLKANVTSENMPRTNKTET 900
917 QLELPVYAVYVMVTSYGVSTKYLNTASNTSRVMOHQVSNLQORSIPISLVLVPV 976
901 QLELPVYAVYVMVTSYGVSTKYLNTASNTSRVMOHQVSNLQORSIPISLVLVPV 960
977 RLNGTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCISIAVCQRIQCDIP 1036
961 RLNGTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCISIAVCQRIQCDIP 1020
1037 FFGIQEBFNATLKNLSFDWYIKTSHNHLIIVSTAEILFNDVSUPTLLPGQAGVRSQTET 1096
1021 FFGIQEBFNATLKNLSFDWYIKTSHNHLIIVSTAEILFNDVSUPTLLPGQAGVRSQTET 1080
1097 KVEPPEVNPPLIIVGSSVGLLALLIITAALYKLGFFKQYKDMGSEGGPPGAEPQ 1153
1081 KVEPPEVNPPLIIVGSSVGLLALLIITAALYKLGFFKQYKDMGSEGGPPGAEPQ 1137

RESULT 10
US-09-902-481A-3
; Sequence 3, Application US/09902481A
; Publication No. US2003005440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY

FILE REFERENCE: A-70586-1/RFT/RMS/RMK
CURRENT APPLICATION NUMBER: US/09/902,481A
CURRENT FILING DATE: 2001-07-09
PRIOR APPLICATION NUMBER: US 60/216,600
PRIOR FILING DATE: 2000-07-07
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn version 3.1
SEQ ID NO 3
LENGTH: 1137
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: synthetic
US-09-902-481A-3

Query Match 98.0%; Score 5839; DB 10; Length 1137;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;
QY 17 FNLDTENAMTFOENARGFGQSVVQLQGSRVVVGAPQEIIVAAANQSGSLVQCDYSTSCSCEPI 76
DB 1 FNLDTENAMTFOENARGFGQSVVQLQGSRVVVGAPQEIIVAAANQSGSLVQCDYSTSCSCEPI 60
QY 77 RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSENTVYKGLCFGLFGSNLRQOPQK 136
DB 61 RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSENTVYKGLCFGLFGSNLRQOPQK 120
QY 137 PPEARLGCQEDSDIAFLIDGSGSIIPHDPRFKKGFVSTVMEQLKKSKTLPISLMQVSESP 196
DB 121 PPEARLGCQEDSDIAFLIDGSGSIIPHDPRFKKGFVSTVMEQLKKSKTLPISLMQVSESP 180
QY 197 RIHFTFEQWNNPRSLIKPIITQLLGRTHATGLRVVRELFTNITNGARKNAFKILFLL 256
DB 181 RIHFTFEQWNNPRSLIKPIITQLLGRTHATGLRVVRELFTNITNGARKNAFKILFLL 240
QY 257 TDGEKFGDLYEDVPEADREGVIRVVGDAFSEKSEKROELNTTASPRDHVFOVN 316
DB 241 TDGEKFGDLYEDVPEADREGVIRVVGDAFSEKSEKROELNTTASPRDHVFOVN 300
QY 317 NFELAKTIQNLREKIFAIETGOTGSSSSFEHEMSQEGFSAALITNGPLLSVGSYDWAG 376
DB 301 NFELAKTIQNLREKIFAIETGOTGSSSSFEHEMSQEGFSAALITNGPLLSVGSYDWAG 360
QY 377 GVFLYTSKESKSTFNNTRVDSMDNDAYLGAALILNRVQSLVGLGAPRYCHIGLVAMFR 436
DB 361 GVFLYTSKESKSTFNNTRVDSMDNDAYLGAALILNRVQSLVGLGAPRYCHIGLVAMFR 420
QY 437 QNTGMWESNANVKGTQICAYFGASCLSDVDVDSNGSTDLVLIGAPHYETRGQVSCPL 496
DB 421 QNTGMWESNANVKGTQICAYFGASCLSDVDVDSNGSTDLVLIGAPHYETRGQVSCPL 480
QY 497 PRGQARWQCDVLYGEGQPGWGFAGALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556
DB 481 PRGQARWQCDVLYGEGQPGWGFAGALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
QY 557 HGTSGSGISPSHSORIASKLSPLQYFGOSLSGGQDLTMDGLVDTLVGAQGHVLLRSQ 616
DB 541 HGTSGSGISPSHSORIASKLSPLQYFGOSLSGGQDLTMDGLVDTLVGAQGHVLLRSQ 600
QY 617 PVLRVKAIMFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLRBEQIQSVVT 676
DB 601 PVLRVKAIMFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLRBEQIQSVVT 660
QY 677 YDLALDSGRPHSRAVFNETKNSRQTVGLTQTCETLKLQPNCEIDPVSPIVLRNLF 736
DB 661 YDLALDSGRPHSRAVFNETKNSRQTVGLTQTCETLKLQPNCEIDPVSPIVLRNLF 720
QY 737 SLVGTPLSAFNGLRPVLAEDAQRLLFTALFPPEKNCNDNITCQDDLSTTFPSMGLDCLVWG 796
DB 721 SLVGTPLSAFNGLRPVLAEDAQRLLFTALFPPEKNCNDNITCQDDLSTTFPSMGLDCLVWG 780
QY 797 GPREFNVTVVRNDGEDSYRTQVTFPEPLDLSYKVSFLQNSQSRWRLACSSASTEV 856
DB 781 GPREFNVTVVRNDGEDSYRTQVTFPEPLDLSYKVSFLQNSQSRWRLACSSASTEV 840

Db 781 GPRENVTVVRNDEGDSYRQVTFPPDLDSYKVKVSTLQORSQSRWELACESASSTEV 840
Qy 857 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEP 916
Db 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEP 900
Qy 917 QLELPVKYAVVMTVSHGVSTKYLNTASENTSRVMOHQYOVSNLQORSLSPLSLVFLVPV 976
Db 901 QLELPVKYAVVMTVSHGVSTKYLNTASENTSRVMOHQYOVSNLQORSLSPLSLVFLVPV 960
Qy 977 RLNQTVIWDPRQVTFSENLSSTCHTKERLPKSHDPLABELKAPVNVCSIAVCQRIQCDDIP 1036
Db 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPKSHDPLABELKAPVNVCSIAVCQRIQCDDIP 1020
Qy 1037 PFGIOSEPNATLGNLSDWYIKTSHNHLIVSTAEILLPNDVFTLLPQCGAFVRSQDET 1096
Db 1021 PFGIOSEPNATLGNLSDWYIKTSHNHLIVSTAEILLPNDVFTLLPQCGAFVRSQDET 1080
Qy 1097 KYBPPEVNPFLIYGVSSVGGLLALLALITAAALYKLGFPKQYKDMWSEGGPPGABPQ 1153
Db 1081 KYBPPEVNPFLIYGVSSVGGLLALLALITAAALYKLGFPKQYKDMWSEGGPPGABPQ 1137

RESULT 11
US-10-116-275-204
; Sequence 204, Application US/10116275
; Publication No. US20030211476A1
; GENERAL INFORMATION:
; APPLICANT: Elan Pharmaceutical Technology
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Brayden, David
; APPLICANT: Byrne, Daragh
; APPLICANT: Lambkin, Imelda
; APPLICANT: Higgins, Lisa
; TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and
; FILE REFERENCE: 51067/20087
; CURRENT APPLICATION NUMBER: US/10/116,275
; NUMBER OF SEQ ID NOS: 349
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 204
; LENGTH: 1163
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-116-275-204

Query Match 58.8%; Score 3503; DB 15; Length 1163;
Best Local Similarity 61.3%; Pred. No. 1.1e-312;
Matches 700; Conservative 138; Mismatches 297; Indels 6; Gaps 4;

Qy 5 VLLLTALTICHGPNLDTENAMTFOENARGFGQSVVQLQSRVWVGAPQEIIVANQORGSLY 64
Db 8 LLLFTALATSLGFLNLTETLAFRVDSDGFGSDVWQYANSWVWVGAPQKITAANQTGGLY 67
Qy 65 QCDYSGSCPELRLOVPEAVNNSLGLSLAATTSPPOLLACGPTVHOTCSENTYVKGLCF 124
Db 68 QCGYSTGACEPIGLQVPEAVNNSLGLSLAATTSPPOLLACGPTVHHECGRNWYLTGLCF 127
Qy 125 LFGSNLRQQPKPEALRCQEDSDIAFLIDCGSGIIPHDFRMEKFVSTVWBLKSKX 184
Db 128 LLGPT--QLTQLPVRQBCPCQEQDIVELIDGSGISSRNFATMNFVRAVISQFQRES 185
Qy 185 TLPSLMQYSEEPRIHTEKFONNPNRSLVKPITOLLGRTHATGIRKVRVRELFINING 244
Db 186 TQPSLMQFNKPTHTTFEPRTSNPLSLASVHQLQFTTATATQNVHRLPHASVG 245
Qy 245 ARKNAFKILVITDGBKDPGLGYEDVPEADREGVIRYVIGVDAPRSEKROELNTIA 304
Db 246 ARSDATKILLIVTDGKKGSDLYDKDVPIMADAAGIIRYAIGVGLAFQNRNSWKLNDIA 305
Qy 305 SKPDRPHVQVNNFEALKTIONQLREKIATIEGTQVSGSSSEHEMSQEGFAALITNGP 364

Db 306 SKPSQEHIFKVEDFDALKDIONQLKEKI PAI EGTETTTSSSSFELEMAQEGFSAVFTDGP 365
Qy 365 LLSTVSGYDWAGGVFLITSKEKSTFPIINNTVDSMDNDAYLGYAAAIIILNRVQSLVLTGAP 424
Db 366 VLGAAGSFTMSGGAFLYPPNNSPTFIMNSQENVDNRSDYILGYSTELAKWGVQSLVLTGAP 425
Qy 425 RYOHIGILVANFRONTGWESNANVKGFOICAYFGASLCSVDVDSNGSTDLVLICAPHYIE 484
Db 426 RYQHTGAVLFTQVSRQWRMKAEVTQTOIGSYFGASLCSVDVDTDGGSTDLVLICAPHYIE 485
Qy 485 QTRGGQVSVCLPRGQBARWQCDAYLIGEQQPAGRFGAALTVLGDVNGDKLTDVAITGAP 544
Db 486 QTRGGQVSVCLPRGQBARWQCDAYLIGEQQPAGRFGAALTVLGDVNGDKLTDVAITGAP 544
Qy 545 GEENRGAIVLPHGTSGSGLSPSHSQRIAGSKLSPRLQYFQSQSLSCGQDLTMDGLVLT 604
Db 545 GEENRGAIVLPHGVLPGLSPSHSQRIAGSKLSPRLQYFQSQSLSCGQDLTMDGLVLT 604
Qy 605 GAQGHVLLLRSQPVLRVKAIMEFNPREVANNVPECNDQVVKGKEAGREVRVCLHVKQSTRD 664
Db 605 GARGQVLLLRTPVLMVGVSMQFIPABIPRSFAFECREQVWSEQLTVQSNICLYIDKESKN 664
Qy 665 RLREGIOQSVTVYDLALDSCRPHSRVFNFTNSTRTQTOVLGTLTQTCETLKLQLPNCIE 724
Db 665 LLGSRDIQSSVTLDLALDPGLSPRATFOHTKRSLSRVRLGLKANCENFLLPSCVE 724
Qy 725 DPVSPVILRLNFSLVGTPLSAFGNLRPLVLAEDAQRLFTALPPFPKNCNDNICODDLSIT 784
Db 725 DSVTPITLRLNFTLVGKPLLAFLNLRPMLAALAQRYFTASLPFEKNCADHICQDNLGIS 784
Qy 785 FSEMSLQVLAVGGPREFNVTVTENDGEDSVRTQVTFPPDLDSYKVKVSTLQORSQSRW 844
Db 785 FSPFGLKSLLVGSLNLENAEVMVWNGEDSVTITTFHPAGLSYRIVAEQKQQLASL 844
Qy 845 RLACESASSTEVSGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVT 904
Db 845 HLTCDASAPVG--SQSTWSTSCRIINHLIFRGAQITFLATPDVSPKAVLGDRLILTANVSS 902
Qy 905 ENNMPTNKTEFQLELPVKYAVVMTVSHGVSTKYLNTAS--ENTSRVMOHQYOVSNLQ 963
Db 903 ENNMPTNKTEFQLELPVKYAVVMTVSHGVSTKYLNTAS--ENTSRVMOHQYOVSNLQ 962
Qy 964 RSLPISLVFLVPLRLNQTVIWDPRQVTFSENLSSTCHTKERLPKSHDPLABELKAPVNV 1023
Db 963 RDLPVSNFVLPVNLQEAVMVDVSHPPQNPSPRLCSSEKIAPPASDPLAHIQNPFVLD 1022
Qy 1024 SIACVQRIQCDDIPFGIOSEPNATLGNLSDWYIKTSHNHLIVSTAEILLPNDVFTLL 1083
Db 1023 SIACVQRIQCDDIPFGIOSEPNATLGNLSDWYIKTSHNHLIVSTAEILLPNDVFTLL 1082
Qy 1084 PQCGAFVRSQDETETKVEPPEVNPFLIYGVSSVGGLLALLALITAAALYKLGFPKQYKDM 1143
Db 1083 PQCGAFVRSQDETETKVEPPEVNPFLIYGVSSVGGLLALLALITAAALYKLGFPKQYKDM 1142
Qy 1144 E 1144
Db 1143 E 1143

RESULT 12
US-09-350-259-4
; Sequence 4, Application US/09350259
; Patent No. US20020062008A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20020062008A1 Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350,259
; CURRENT FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: 09/193,043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497

```

; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1163
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-350-259-4

Query Match      58.4%; Score 3480; DB 9; Length 1163;
Best Local Similarity 61.1%; Pred. No. 1.5e-310;
Matches 697; Conservative 144; Mismatches 294; Indels 6; Gaps 4;

QY 5 VLLLTALTCHGNLDTENAMTFOENARGFGQSVVQLQGSRRVVGAPQBEIVAAHQSGSLY 64
DB 8 LLLLTALATSLGFLNLTDELTAPRVDSAGFGDSVQVYANSMVVGAPQKIILAAQIGGLY 67

QY 65 QCDYSGTSCPIRLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSENTYVKGLCF 124
DB 68 QCGYSTGACPIGLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSENTYVKGLCF 127

QY 125 LFGSNLRQOQPKFPPALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKEFVSTVMEQKKSK 184
DB 128 LAGPT--QLTQRLPVSRQCPQEQDIVFLIDGSGSISSRNFAFMNFVRAVISQFRPS 185

QY 185 TLFPSLMQVSEBRIHTFKFQNNPNRSLVKPITQLGRHTATGIRKVVRELFNITNG 244
DB 186 TQFSLMQVSNKPTQHTFEBFRTSNPLSLASVHQLQGFYTTATQVHRLPHASVYG 245

; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1163
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-350-259-4

Query Match      58.4%; Score 3480; DB 9; Length 1163;
Best Local Similarity 61.1%; Pred. No. 1.5e-310;
Matches 697; Conservative 144; Mismatches 294; Indels 6; Gaps 4;

QY 785 FSPFGLKSLVGSNLELNAEYVWMDGSDSYGTTITFSGHPAGLSTRYVARGQKQLRSL 844
QY 845 RLACESASSTVSGALKSTSCSINHPIPPENSEVTNITFDVDSKASLGNKLLKANVTS 904
DB 845 HLTC--CSAPVGSQQTSTSCSINHPIPRGAQITFLATFDVSPVAVGLDRLLIANVSS 902
QY 905 ENNMPTNKTBFQLELPVKYAVYVWVTSKYSTKYNLTAS--ENTSRVWQHOYQVSNLQ 963
DB 903 ENNIPRTSKTIFQLELPVKYAVYVWVTSKYSTKYNLTAS--ENTSRVWQHOYQVSNLQ 962
QY 964 RSLPISLVLPLVRLNQTWINDRPOVTFSENLSSTCHTKERLPSPHSDFLAELRKAPVNC 1023
DB 963 RDLFVSNFVWVPELNAQAVWMDVEVSHPOPSLSCSEKTAAPPASDFLAHIQKQPVLC 1022
QY 1024 STAVCORIQCDIPFGIOEENFATLGNLSPDWYIKTSHNHLIIVSTAEILPNDVFTLL 1083
DB 1023 STAGCLRERCDVPSPVQBELDFTLGNLSPDWYIKTSHNHLIIVSTAEILPNDVFTLL 1082
QY 1084 PGOCAFVSQSTETKVEPEVFNPLIIVGSSVGGILLALITAAALYKLGPPKROVKDMS 1143
DB 1083 PGOCAFVSQSTETKVEPEVFNPLIIVGSSVGGILLALITAAALYKLGPPKROVKDMS 1142
QY 1144 E 1144
DB 1143 E 1143

RESULT 13
US-09-891-943-4
; Sequence 4, Application US/09891943
; Publication No. US20030077278A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: NO. US20030077278A1 Human 2
; FILE REFERENCE: 27866735004
; CURRENT APPLICATION NUMBER: US/09/891,943
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1163
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-891-943-4

Query Match      58.4%; Score 3480; DB 10; Length 1163;
Best Local Similarity 61.1%; Pred. No. 1.5e-310;
Matches 697; Conservative 144; Mismatches 294; Indels 6; Gaps 4;

QY 5 VLLLTALTCHGNLDTENAMTFOENARGFGQSVVQLQGSRRVVGAPQBEIVAAHQSGSLY 64
DB 8 LLLLTALATSLGFLNLTDELTAPRVDSAGFGDSVQVYANSMVVGAPQKIILAAQIGGLY 67

QY 65 QCDYSGTSCPIRLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSENTYVKGLCF 124
DB 68 QCGYSTGACPIGLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSENTYVKGLCF 127

QY 125 LFGSNLRQOQPKFPPALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKEFVSTVMEQKKSK 184
DB 128 LAGPT--QLTQRLPVSRQCPQEQDIVFLIDGSGSISSRNFAFMNFVRAVISQFRPS 185

QY 185 TLFPSLMQVSEBRIHTFKFQNNPNRSLVKPITQLGRHTATGIRKVVRELFNITNG 244
DB 186 TQFSLMQVSNKPTQHTFEBFRTSNPLSLASVHQLQGFYTTATQVHRLPHASVYG 245

```


QY 725 DVSPIVLRNLSFVLGTFSAFAGNLRPVLAEDAQRLEFALFPPEKNCNDNIQODLSIT 784
DB 722 DVVSPILHLNLSFVLRSPFSPQNLRPVLA VGSQDLFTASLPPEKNCQDGLCEGDLGVT 781
QY 785 FSPMSLCLVVGGRPREFNVTVVNDGDSYRTOVTFPPPLDLSYKXVSTLQORQORSW 844
DB 782 LSPSGLOTLTVGSSLELNVIIVTMNAGEDSYGTVVSUYIPAGLSHRVSGAQOPHQSAL 841
QY 845 RLACESASSTEVSGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTS 904
DB 842 RLACETV-PTDEG-LASSRCSVNHPIFHEGSGNGTFTVTFDVSYSKATLGRMLWRASASS 899
QY 905 ENNMPRTNKTPEQLELPVKYAVMVVSHGVSTKYLNFP-TASENTSRVMQHOYQVSNLQ 963
DB 900 ENNKASSSKATFQLELPVKYAVMTMISROEBESTKYFNFAYSDEKMKAEHRVNNLSQ 959
QY 964 RSLPISLVLVPLVRLNQTVDWRPQVTFSENLSTCHTKERLPSHSDFLAELKAPVNC 1023
DB 960 RDLAISINFWPVLNGVAVDVMWVMEAPQSLL-PCVSEKPKQHSDFLTQISRPMLDC 1017
QY 1024 SIACQRIQCDIPFGIOBEFNATLKNLSFDWIKTSHHLLIVSTABILFNDSVFTLL 1083
DB 1018 SIADCLQFRCDVPSFVSQEBLDFTLKGNLSFGVRETLQKKVLVSVVAEITFDTSVYSQL 1077
QY 1084 PQCAFVRVRSOTETKVEPPEFVDPNPLVGVSSVGLLILLALITAAALYKLGFFKQYKDMS 1143
DB 1078 PQCAFVRVRSOTETKVEPPEFVDPNPLVGVSSVGLLILLALITATLYKLGFFKQYKEMLE 1137
QY 1144 E 1144
DB 1138 D 1138

RESULT 15

US-09-891-943-2
; Sequence 2, Application US/09891943
; Publication No. US20030077278A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; TITLE OF INVENTION: No. US20030077278A1 Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/891,943
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1161
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-891-943-2

Query Match 58.0%; Score 3455; DB 10; Length 1161;
Best Local Similarity 59.9%; Pred. No. 2.9e-308;
Matches 684; Conservative 162; Mismatches 287; Indels 8; Gaps 6;
QY 5 VLLTALTLCHGFNLDENAMTFQENARFGQSVVQLQGRVVGAPQOEVANQORGLY 64
DB 5 VLLLSVLASVHGFLNLDVEEPTIFQEDAGGQSVVQFGSRLVVGAPLEVVAANQTRLY 64
QY 65 QCDVSTGSCPEIRLOVPVAVNMSIGLSLAATSPPOLLAGPTVHOTCSNTVVKGLCF 124
DB 65 DCAATGACQPIPLHRIPEAVNMSIGLTLASTNGSRLACGPTLHRCVGENSYSGKSL 124

QY 125 LFGSNLRQOPQKFPALRGCCPOEDSDIAPLIDGSGSIIPHDPRMKBFVSTWEOQLKSK 184
DB 125 LLGSRW-ELIQVTPDAPTECPHQEMDIVPLIDGSGSIDQDNFQKGGFVQAVMGQFEGD 183
QY 185 TLFSLMQVSEEPRIHFTFKFQNNPNRSLVKPIQTLQCRHTATCIRKVVARELFNITNG 244
DB 184 TLFALMQVSNLKLHFTFTQFATSPSQSLVDPIVQLKGLTATACILTAVVQLFHHKG 243
QY 245 ARNAPKILVITDKEFGDPLGYEDVTPEDRDEGVIRVIGVDAFRSEKRSQELNTIA 304
DB 244 ARKSARKILVITDQKQVDPLEYSVTPQAEKAGIIRYAIAGVGHAFQFTARQELNTIS 303
QY 305 SPPPDHVPQVNNFALKTIONLREKIFAIGTGTGSSSSFEHEMSQGFSAATNSGP 364
DB 304 SAPPQHVFKVNFNFAALGSIQOLQEKIYAVSGTQSRASSSQHEMSQGFSTALTMGL 363
QY 365 LLSVGSVDWAGGVFLYTSKEKSTFINTRVDSMDNDAYLYGAAAIIILNRVQSLVIGAP 424
DB 364 FLGAVGSPSWSGAFLYPPNMSPTFINMSQENVMDRDSYLGYSTELALWKGQNLVIGAP 423
QY 425 RYCHICLVAMPQNTOMWESNANVGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYTE 484
DB 424 RYCHTGKAVITQVGRQWRKKAETGTQIGSTFGASLCSVDVDSGSDTLILIGAPHYTE 483
QY 485 QTRGGQVSVCPFRGORARQCDVLYGSGQGPWCRFGAALTVLGDVNGDKLTDVAIGAP 544
DB 484 QTRGGQVSVCPFRGORVQVQCDVLRGQBQHPWGRFGAALTVLGDVNGDKLTDVAIGAP 543
QY 545 GEDNRGAVLPHGTSGSGISPSHSQRIAGSKLSRLOVFGOSLGGQDLTMDGLVDLTV 604
DB 544 GECENRGAVLPHGASEGISPSHSQRIAGSKLSRLOVFGOSLGGQDLTMDGLMDLAV 603
QY 605 GAQGHVLLRSQVLRVKAIMEFNPREVARNYFECNDQVVGKBEAGEVRLHVQKSTRD 664
DB 604 GARGQVLLRLSLPVLKVGVMRFSPEVAKAVYRQWEEKPSALEAGDATVCLTIQKSLD 663
QY 665 RLREGIOQSVVTVYDIALDSGRPHSAVNETKNSRTOQVGLTCTCETLKLQPNCLIE 724
DB 664 QL--GDIQSSVRPDLALDDEGRUTSAIFNETKNPTLTRKTLGLGHCSTLKLLEPCVE 721
QY 725 DVSPIVLRNLSFVLGTFSAFAGNLRPVLAEDAQRLEFALFPPEKNCNDNIQODLSIT 784
DB 722 DVVSPILHLNLSFVLRSPFSPQNLRPVLA VGSQDLFTASLPPEKNCQDGLCEGDLGVT 781
QY 785 FSPMSLCLVVGGRPREFNVTVVNDGDSYRTOVTFPPPLDLSYKXVSTLQORQORSW 844
DB 782 LSPSGLOTLTVGSSLELNVIIVTMNAGEDSYGTVVSUYIPAGLSHRVSGAQOPHQSAL 841
QY 845 RLACESASSTEVSGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTS 904
DB 842 RLACETV-PTDEG-LASSRCSVNHPIFHEGSGNGTFTVTFDVSYSKATLGRMLWRASASS 899
QY 905 ENNMPRTNKTPEQLELPVKYAVMVVSHGVSTKYLNFP-TASENTSRVMQHOYQVSNLQ 963
DB 900 ENNKASSSKATFQLELPVKYAVMTMISROEBESTKYFNFAYSDEKMKAEHRVNNLSQ 959
QY 964 RSLPISLVLVPLVRLNQTVDWRPQVTFSENLSTCHTKERLPSHSDFLAELKAPVNC 1023
DB 960 RDLAISINFWPVLNGVAVDVMWVMEAPQSLL-PCVSEKPKQHSDFLTQISRPMLDC 1017
QY 1024 SIACQRIQCDIPFGIOBEFNATLKNLSFDWIKTSHHLLIVSTABILFNDSVFTLL 1083
DB 1018 SIADCLQFRCDVPSFVSQEBLDFTLKGNLSFGVRETLQKKVLVSVVAEITFDTSVYSQL 1077
QY 1084 PQCAFVRVRSOTETKVEPPEFVDPNPLVGVSSVGLLILLALITAAALYKLGFFKQYKDMS 1143
DB 1078 PQCAFVRVRSOTETKVEPPEFVDPNPLVGVSSVGLLILLALITATLYKLGFFKQYKEMLE 1137
QY 1144 E 1144
DB 1138 D 1138

us-09-902-481b-1.rapb

Tue Jun 8 05:59:06 2004

Search completed: June 7, 2004, 17:38:40
Job time : 43.4378 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 17:07:20 ; Search time 18.6066 Seconds
(without alignments)
3199.127 Million cell updates/sec

Title: US-09-902-481b-1
Perfect score: 5956
Sequence: 1 MALRVLLLTALTLCGFNLD.....FKRQYKDMSEGGPPGAEPRQ 1153

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:**
1: /cgn2_6/ptodata/2/iaa/5A COMB.pdp.*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pdp.*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pdp.*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pdp.*
5: /cgn2_6/ptodata/2/iaa/PCTUS COMB.pdp.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pdp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5956	100.0	1153	1	US-08-173-497-3
2	5956	100.0	1153	1	US-08-286-889-3
3	5956	100.0	1153	1	US-08-485-618-3
4	5956	100.0	1153	1	US-08-362-652-3
5	5956	100.0	1153	2	US-08-605-672-3
6	5956	100.0	1153	2	US-08-482-293A-3
7	5956	100.0	1153	2	US-08-943-363-3
8	5956	100.0	1153	3	US-09-193-043-3
9	5956	100.0	1153	4	US-09-688-307A-3
10	5956	100.0	1153	4	US-09-350-259-3
11	5925.5	99.5	1152	2	US-08-476-062A-43
12	5925.5	99.5	1152	6	PCT-US96-01314-43
13	5925.5	99.5	1152	6	PCT-US96-01314-43
14	3503	58.8	1163	5	US-08-476-062A-44
15	3503	58.8	1163	5	PCT-US96-01314-44
16	3480	58.4	1163	1	US-08-173-497-4
17	3480	58.4	1163	1	US-08-286-889-4
18	3480	58.4	1163	1	US-08-485-618-4
19	3480	58.4	1163	1	US-08-362-652-4
20	3480	58.4	1163	2	US-08-605-672-4
21	3480	58.4	1163	2	US-08-482-293A-4
22	3480	58.4	1163	2	US-08-943-363-4
23	3480	58.4	1163	3	US-09-193-043-4
24	3480	58.4	1163	4	US-09-688-307A-4
25	3480	58.4	1163	4	US-09-350-259-4
26	3455	58.0	1161	1	US-08-173-497-2
27	3455	58.0	1161	1	US-08-286-889-2

US-08-173-497-3
Sequence 3, Application US/08173497
Patent No. 5437958
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
APPLICANT: Van Der Vieren, Monica
TITLE OF INVENTION: No. 5437958e1 Human 2 Integrin Alpha
TITLE OF INVENTION: Subunit
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 S. Wacker Drive, 6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606-6402
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/173,497
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 5437958and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 27866/31363
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1153 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-173-497-3

Query Match 100.0%; Score 5956; DB 1; Length 1153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 0; Gaps 0;

QY 1 MALRVLLLTALTLCGFNLDTENAMTFQENARGGQSVVQLQSGRVVVGAPQEIIVAAKOR 60
DB 1 MALRVLLLTALTLCGFNLDTENAMTFQENARGGQSVVQLQSGRVVVGAPQEIIVAAKOR 60

ALIGNMENTS

RESULT 1
US-08-173-497-3
Sequence 3, Application US/08173497
Patent No. 5437958
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
APPLICANT: Van Der Vieren, Monica
TITLE OF INVENTION: No. 5437958e1 Human 2 Integrin Alpha
TITLE OF INVENTION: Subunit
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 S. Wacker Drive, 6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606-6402
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/173,497
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 5437958and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 27866/31363
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1153 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-173-497-3

Query Match 100.0%; Score 5956; DB 1; Length 1153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 0; Gaps 0;

QY 1 MALRVLLLTALTLCGFNLDTENAMTFQENARGGQSVVQLQSGRVVVGAPQEIIVAAKOR 60
DB 1 MALRVLLLTALTLCGFNLDTENAMTFQENARGGQSVVQLQSGRVVVGAPQEIIVAAKOR 60

ALIGNMENTS

QY 61 GSLYCDYSTGSCPIRLQVPEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYK 120
DB 61 GSLYCDYSTGSCPIRLQVPEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYK 120
QY 121 GLCFPGNLRQOQKPPPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEFPVSTVMEQL 180
DB 121 GLCFPGNLRQOQKPPPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEFPVSTVMEQL 180
QY 181 KSKTLPISLMQYSEEFRIHFTFKFQNNPNRSLVKPITOLLGRTHRTATGIRKVVRELFN 240
DB 181 KSKTLPISLMQYSEEFRIHFTFKFQNNPNRSLVKPITOLLGRTHRTATGIRKVVRELFN 240
QY 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVIPADREGVIRYVIGVDAPRSEKSRQEL 300
DB 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVIPADREGVIRYVIGVDAPRSEKSRQEL 300
QY 301 NTIASKPRDHVQVNNFEALKTIONLREKIPALBGTQGGSSSFHEMSQSGFSAIT 360
DB 301 NTIASKPRDHVQVNNFEALKTIONLREKIPALBGTQGGSSSFHEMSQSGFSAIT 360
QY 361 SNGPSTVGSYDWDAGGVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLV 420
DB 361 SNGPSTVGSYDWDAGGVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLV 420
QY 421 LGABRYOHIGLVAMFRONTGMESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAP 480
DB 421 LGABRYOHIGLVAMFRONTGMESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAP 480
QY 481 HYETQTRGGQVSVCPFRGQARQCDVLYGEOQPGWRFGAALTVLGVDVNGDKLTDVA 540
DB 481 HYETQTRGGQVSVCPFRGQARQCDVLYGEOQPGWRFGAALTVLGVDVNGDKLTDVA 540
QY 541 IGAPGEDNRGANVLYHGTSGSGISPSHSORIASGKLSPLRYPGQSLGGQDLTMDGLV 600
DB 541 IGAPGEDNRGANVLYHGTSGSGISPSHSORIASGKLSPLRYPGQSLGGQDLTMDGLV 600
QY 601 DLTGACQGHVLLRSQVLRVKAIMEFNPREVARNFECNDQVVKGEAGEVAVCLHVQK 660
DB 601 DLTGACQGHVLLRSQVLRVKAIMEFNPREVARNFECNDQVVKGEAGEVAVCLHVQK 660
QY 720 STDRRLREGIQSVVYTDLALDGRPHSAVFNETKNSRTRQTVGLGLTQCTETKLQLP 720
DB 720 STDRRLREGIQSVVYTDLALDGRPHSAVFNETKNSRTRQTVGLGLTQCTETKLQLP 720
QY 721 NCIEDPVSPIVLRNLSVGLTSAFGLNLPVLAEDAQRFTALPPEKNCNDNICQDD 780
DB 721 NCIEDPVSPIVLRNLSVGLTSAFGLNLPVLAEDAQRFTALPPEKNCNDNICQDD 780
QY 781 LSITFSFMSLDCLVVGCPREFNVTVTVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNR 840
DB 781 LSITFSFMSLDCLVVGCPREFNVTVTVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNR 840
QY 841 QRSRLACESASSTEVSGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKA 900
DB 841 QRSRLACESASSTEVSGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKA 900
QY 901 NVTSENNMPRTNTEFQLELPVKYAVTVVMTSHGVSTKYLNFTASENTSRVMQHQVSN 960
DB 901 NVTSENNMPRTNTEFQLELPVKYAVTVVMTSHGVSTKYLNFTASENTSRVMQHQVSN 960
QY 961 LGQSLDPSILVFLVPLNQTIVDRQVTFSENLSSTCTKRELPSHSDFTLAEKAPV 1020
DB 961 LGQSLDPSILVFLVPLNQTIVDRQVTFSENLSSTCTKRELPSHSDFTLAEKAPV 1020
QY 1021 VNCISIAVCQRIQCDIPFFQIEQEFNATLKGNLSPDWYIKTSHNLLITVSTAEILFNDVSF 1080
DB 1021 VNCISIAVCQRIQCDIPFFQIEQEFNATLKGNLSPDWYIKTSHNLLITVSTAEILFNDVSF 1080
QY 1081 TLLPGQAFVRSQETETKVPFEPVNPPLIVGSSVGLLALLIATAALYKLGFFKQYKD 1140
DB 1081 TLLPGQAFVRSQETETKVPFEPVNPPLIVGSSVGLLALLIATAALYKLGFFKQYKD 1140

QY 1141 MMSEGGPPGRNEPQ 1153
DB 1141 MMSEGGPPGRNEPQ 1153
RESULT 2
US-08-286-889-3
; Sequence 3, Application US/08286889
; Patent No. 5470953
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Mich
; TITLE OF INVENTION: No. 5470953el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/286,889
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: P38,659
; REFERENCE/DOCKET NUMBER: 27866/32168
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-286-889-3
Query Match 100.0%; Score 5956; DB 1; Length 1153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALRVLLLTALTLCHGFNLDTENAMTFQENARGFGQSVVQLQSGRVVVGAPQEI VAAANOR 60
DB 1 MALRVLLLTALTLCHGFNLDTENAMTFQENARGFGQSVVQLQSGRVVVGAPQEI VAAANOR 60
QY 61 GSLYCDYSTGSCPIRLQVPEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYK 120
DB 61 GSLYCDYSTGSCPIRLQVPEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYK 120
QY 121 GLCFPGNLRQOQKPPPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEFPVSTVMEQL 180
DB 121 GLCFPGNLRQOQKPPPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEFPVSTVMEQL 180
QY 181 KSKTLPISLMQYSEEFRIHFTFKFQNNPNRSLVKPITOLLGRTHRTATGIRKVVRELFN 240
DB 181 KSKTLPISLMQYSEEFRIHFTFKFQNNPNRSLVKPITOLLGRTHRTATGIRKVVRELFN 240
QY 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVIPADREGVIRYVIGVDAPRSEKSRQEL 300
DB 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVIPADREGVIRYVIGVDAPRSEKSRQEL 300

301 NTIASKPPDRHVQVNNFEALKTIONQREKIFAIEGTOTGSSSSFEHMSQSGPSAAIT 360
301 NTIASKPPDRHVQVNNFEALKTIONQREKIFAIEGTOTGSSSSFEHMSQSGPSAAIT 360
361 SNGPLLSVGSYDWAGVFLYTSKSTFTINTRVDSMDNDAYLGVAIAIILNRVQSIV 420
361 SNGPLLSVGSYDWAGVFLYTSKSTFTINTRVDSMDNDAYLGVAIAIILNRVQSIV 420
421 LGAPRYQHIGLVAMFRQNTGMSNANVKGTOIGVFGASLCSVDVDSNGSDTLVLIGAP 480
421 LGAPRYQHIGLVAMFRQNTGMSNANVKGTOIGVFGASLCSVDVDSNGSDTLVLIGAP 480
481 HYTEQTRGGQVSVCPILPRGORARWQCDVLYGBOGQWGRFGAALTVLGSDVNGDKLTDA 540
481 HYTEQTRGGQVSVCPILPRGORARWQCDVLYGBOGQWGRFGAALTVLGSDVNGDKLTDA 540
541 IGAPGEDNRGAYLPHGTSGSGISPSHSORIASKLSRLQVFGOSLSGGQDLTWGSLV 600
541 IGAPGEDNRGAYLPHGTSGSGISPSHSORIASKLSRLQVFGOSLSGGQDLTWGSLV 600
601 DLTVGAGQHVLILRSQPLVLRVKAIMEFNPFEVARNVPECNDQVVGKEAGEVRVCLHVQK 660
601 DLTVGAGQHVLILRSQPLVLRVKAIMEFNPFEVARNVPECNDQVVGKEAGEVRVCLHVQK 660
661 STDRLRREGIOISVTVYDIALDSGRPHSRVAFNETKNSRRQTVGLTQTCTETLKLQLP 720
661 STDRLRREGIOISVTVYDIALDSGRPHSRVAFNETKNSRRQTVGLTQTCTETLKLQLP 720
721 NCIEDPVSPIVLRNLSVLTGTPLSAFCNLRPVLAEDAQRLFTALFPFKKNCNDNICQDD 780
721 NCIEDPVSPIVLRNLSVLTGTPLSAFCNLRPVLAEDAQRLFTALFPFKKNCNDNICQDD 780
781 LSITPFSMLDCLVWGPREFNVTVVRNDGDSYRTQVTPFPPLDLSYRKVSTLQNS 840
781 LSITPFSMLDCLVWGPREFNVTVVRNDGDSYRTQVTPFPPLDLSYRKVSTLQNS 840
841 ORSWRLACSSASTVSGALKSTCSINHPPIFENSEVTFTNITFDVDSKASLGNKLLKA 900
841 ORSWRLACSSASTVSGALKSTCSINHPPIFENSEVTFTNITFDVDSKASLGNKLLKA 900
901 NVTSENMPRTNKTETQLELPVKYAVVMTVSHGVSTKYLNFTASNTSRVQHOVQSN 960
901 NVTSENMPRTNKTETQLELPVKYAVVMTVSHGVSTKYLNFTASNTSRVQHOVQSN 960
961 LGORSIPISLVFLVRLNQTIVDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPV 1020
961 LGORSIPISLVFLVRLNQTIVDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPV 1020
1021 VNCISIAVCORIOCDIPFFGQIEEFNATLKGNTSFDWYIKTSHNHLIIVSTABILFNDSPV 1080
1021 VNCISIAVCORIOCDIPFFGQIEEFNATLKGNTSFDWYIKTSHNHLIIVSTABILFNDSPV 1080
1081 TLLPGQGAFFRSQTEKVPPEVNPDLPLVGSVGGILLALITAALYKLGFKKQYKD 1140
1081 TLLPGQGAFFRSQTEKVPPEVNPDLPLVGSVGGILLALITAALYKLGFKKQYKD 1140
1141 MNSEGGPPGAEPQ 1153
1141 MNSEGGPPGAEPQ 1153

RESULT 3
US-08-485-618-3
Sequence 3, Application US/08485618
Patent No. 5728533
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
TITLE OF INVENTION: No. 5728533el Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,618
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32797
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1153 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-485-618-3

Query Match 100.0%; Score 5956; DB 1; Length 1153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRVLLLTALTLCGPNLDTENAMTFOENARGFQSVVQLQSGRVVVGAPQEI VAAQR 60
DB 1 MALRVLLLTALTLCGPNLDTENAMTFOENARGFQSVVQLQSGRVVVGAPQEI VAAQR 60
QY 61 GSLYQCDYSTGSCFIRLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVK 120
DB 61 GSLYQCDYSTGSCFIRLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVK 120
QY 121 GLCFLFGSNLRQOPKPFPEARLRCQPEDSDIAFLIDGSGSIIPHDPRKKEFVSTVMSQL 180
DB 121 GLCFLFGSNLRQOPKPFPEARLRCQPEDSDIAFLIDGSGSIIPHDPRKKEFVSTVMSQL 180
QY 181 KKSKTFLSLMQSYSEFRIFHTFKCFQNNPNRSLVKPITQLLGRTHHTATGIRKVVRELPN 240
DB 181 KKSKTFLSLMQSYSEFRIFHTFKCFQNNPNRSLVKPITQLLGRTHHTATGIRKVVRELPN 240
QY 241 ITNGARKNAFKILVWITDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPRSEKSRQL 300
DB 241 ITNGARKNAFKILVWITDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPRSEKSRQL 300
QY 301 NTIASKPPDRHVQVNNFEALKTIONQREKIFAIEGTOTGSSSSFEHMSQSGPSAAIT 360
DB 301 NTIASKPPDRHVQVNNFEALKTIONQREKIFAIEGTOTGSSSSFEHMSQSGPSAAIT 360
QY 361 SNGPLLSVGSYDWAGVFLYTSKSTFTINTRVDSMDNDAYLGVAIAIILNRVQSIV 420
DB 361 SNGPLLSVGSYDWAGVFLYTSKSTFTINTRVDSMDNDAYLGVAIAIILNRVQSIV 420
QY 421 LGAPRYQHIGLVAMFRQNTGMSNANVKGTOIGVFGASLCSVDVDSNGSDTLVLIGAP 480

421 LGAPRYOHIGLVANFRONTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAP 480
481 HYETQTRGGQSVCPPLPRGQARWQCDVLYGEGQGWGFGAALTIVLGDVNGDKLTDVA 540
481 HYETQTRGGQSVCPPLPRGQARWQCDVLYGEGQGWGFGAALTIVLGDVNGDKLTDVA 540
541 IGAPGEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLV 600
541 IGAPGEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLV 600
601 DLTGAGQSHVLLRSQVPLRVKALMEFNPREVARNVPECNDQVVGKEAGEVRVCLHVQK 660
601 DLTGAGQSHVLLRSQVPLRVKALMEFNPREVARNVPECNDQVVGKEAGEVRVCLHVQK 660
661 STRDLREGQIQSVVYDLDLDSRPHSRVAFNFKTSTRQVGLTQTCETLKLQLP 720
661 STRDLREGQIQSVVYDLDLDSRPHSRVAFNFKTSTRQVGLTQTCETLKLQLP 720
721 NCIEDPVPVILRNFLNFTPLSAGNLRPLVLAEDAQRLLFTALFPPEKNCMDNICQDD 780
721 NCIEDPVPVILRNFLNFTPLSAGNLRPLVLAEDAQRLLFTALFPPEKNCMDNICQDD 780
781 LSITPFSMSLDCLVVGPRFNFTVTVRNDEGDSYRTQVTFEPFLDLSYKVSITLQORS 840
781 LSITPFSMSLDCLVVGPRFNFTVTVRNDEGDSYRTQVTFEPFLDLSYKVSITLQORS 840
841 QRSWLACASSTEVSGALKSTSCSINHPIPPENSEVTFTNITPDVDSKASLGNKLLKA 900
841 QRSWLACASSTEVSGALKSTSCSINHPIPPENSEVTFTNITPDVDSKASLGNKLLKA 900
901 NVTSENNMPNKTTEFOLELPVKAVYVMTVSHGVSTKYLNFTASENTSRVMQHOYQSVN 960
901 NVTSENNMPNKTTEFOLELPVKAVYVMTVSHGVSTKYLNFTASENTSRVMQHOYQSVN 960
961 LGQSLPISLVLVPLVRLNQTIVINDRPQVTFPSENLSSTCHTKERLPSPHSDFLAELKAPV 1020
961 LGQSLPISLVLVPLVRLNQTIVINDRPQVTFPSENLSSTCHTKERLPSPHSDFLAELKAPV 1020
1021 VNCISACQRIQCDIPFGIOEBFNATLKNLSFDWIKTSHNHLIVSTAEILFNDVSF 1080
1021 VNCISACQRIQCDIPFGIOEBFNATLKNLSFDWIKTSHNHLIVSTAEILFNDVSF 1080
1081 TLLPGQAFVRSQETKVEPPEVPNPPLIVGSSVGGLLIALLITAAALYKLGFFKQYKD 1140
1081 TLLPGQAFVRSQETKVEPPEVPNPPLIVGSSVGGLLIALLITAAALYKLGFFKQYKD 1140
1141 MMSSEGGPPGABPQ 1153
1141 MMSSEGGPPGABPQ 1153

RESULT 4
US-08-362-652-3
; Sequence 3, Application US/08362652
; Patent No. 5766850
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; TITLE OF INVENTION: No. 5766850el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/362,652
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32391
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1153 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-362-652-3

Query Match 100.0%; Score 5956; DB 1; Length 1153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MALRVLLLTALTLCGHNLDNENAMTFQENARGFGQSVVQLOGSRVVVGAPQEIIVANOR 60
Db 1 MALRVLLLTALTLCGHNLDNENAMTFQENARGFGQSVVQLOGSRVVVGAPQEIIVANOR 60
Qy 61 GSLYQCDYSTGCEPRLQVPEAVNMSLGLSLAATTSPPOLLACQPTVHQCSENTYK 120
Db 61 GSLYQCDYSTGCEPRLQVPEAVNMSLGLSLAATTSPPOLLACQPTVHQCSENTYK 120
Qy 121 GUCFLPGSNLRQOPKPFPEALRGCCQEDSDIAFLIDGSGSIIPHPFRMKCFVSTVMEQL 180
Db 121 GUCFLPGSNLRQOPKPFPEALRGCCQEDSDIAFLIDGSGSIIPHPFRMKCFVSTVMEQL 180
Qy 181 KSKTILFSLMXYSEBFRIFHTPKFQNNPNPSLVKPIITQLLGRTHATGIRKVVRELFN 240
Db 181 KSKTILFSLMXYSEBFRIFHTPKFQNNPNPSLVKPIITQLLGRTHATGIRKVVRELFN 240
Qy 241 ITNGARKNAFKILVITDGEKFGDPIGYEDVPEADREGVIRVVGDAFSEKSRQEL 300
Db 241 ITNGARKNAFKILVITDGEKFGDPIGYEDVPEADREGVIRVVGDAFSEKSRQEL 300
Qy 301 NTIASKPPRDHVPQVNNFEALKTIONQLREKIFALEGTQTGSSSSFEHMSQEGFSAIT 360
Db 301 NTIASKPPRDHVPQVNNFEALKTIONQLREKIFALEGTQTGSSSSFEHMSQEGFSAIT 360
Qy 361 SNGPLSTVSGVDWAGGVLYTSKEKSTEINNTVDSMDNDAYLGYAAAILLRNVQSLV 420
Db 361 SNGPLSTVSGVDWAGGVLYTSKEKSTEINNTVDSMDNDAYLGYAAAILLRNVQSLV 420
Qy 421 LGAPRYOHIGLVANFRONTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAP 480
Db 421 LGAPRYOHIGLVANFRONTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAP 480
Qy 481 HYETQTRGGQSVCPPLPRGQARWQCDVLYGEGQGWGFGAALTIVLGDVNGDKLTDVA 540
Db 481 HYETQTRGGQSVCPPLPRGQARWQCDVLYGEGQGWGFGAALTIVLGDVNGDKLTDVA 540
Qy 541 IGAPGEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLV 600
Db 541 IGAPGEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLV 600
Qy 601 DLTGAGQSHVLLRSQVPLRVKALMEFNPREVARNVPECNDQVVGKEAGEVRVCLHVQK 660
Db 601 DLTGAGQSHVLLRSQVPLRVKALMEFNPREVARNVPECNDQVVGKEAGEVRVCLHVQK 660

601 DLTGAGQGHVLLRSQPVLRVKAIMEFNPRAVARNVFCNDQVVKGEAGEVRVCLHVOK 660
661 STRDLREGQIQSVVYDIALDSGRPHSRAVFNETKSTRQVGLTQTCTETLKLQLP 720
661 STRDLREGQIQSVVYDIALDSGRPHSRAVFNETKSTRQVGLTQTCTETLKLQLP 720
721 NCIEDPVSPIVLRNLSVGTPLSAFGNLRPVLAEDAQRULFTALPFPEKNCNDNICQDD 780
721 NCIEDPVSPIVLRNLSVGTPLSAFGNLRPVLAEDAQRULFTALPFPEKNCNDNICQDD 780
781 LSITFSMSLDCLVGGPREFNTVTVRNDGDSYRTQVTFPPPLDLSYRKVSTLQNR 840
781 LSITFSMSLDCLVGGPREFNTVTVRNDGDSYRTQVTFPPPLDLSYRKVSTLQNR 840
841 QRSWRLACESASSTVSALKSTSCSINHPIPFENSEVFNITFDVDSKASIGNKLLKA 900
841 QRSWRLACESASSTVSALKSTSCSINHPIPFENSEVFNITFDVDSKASIGNKLLKA 900
901 NVTSENMPNTKTEFOLELPVKYAVVMTSHGVSTKYLNFASENTSRVMOHQVSN 960
901 NVTSENMPNTKTEFOLELPVKYAVVMTSHGVSTKYLNFASENTSRVMOHQVSN 960
961 LQORSIPISVLPVRLNQTIVDRPQVTFSENLSSTCHTKERLPKSHSDFLAELRKAPV 1020
961 LQORSIPISVLPVRLNQTIVDRPQVTFSENLSSTCHTKERLPKSHSDFLAELRKAPV 1020
1021 VNCISIAVCORIQCDIPFGIOEBFNATLKGNSFDWYIKTSHNHLIVSTABILFNDVVF 1080
1021 VNCISIAVCORIQCDIPFGIOEBFNATLKGNSFDWYIKTSHNHLIVSTABILFNDVVF 1080
1081 TLLPGQAFVRSQTEKVPFEPVFNPLPLIVGSSVGLLILALITAAALYKLGFFKQYKD 1140
1081 TLLPGQAFVRSQTEKVPFEPVFNPLPLIVGSSVGLLILALITAAALYKLGFFKQYKD 1140
1141 MMSEGGPPGAEPQ 1153
1141 MMSEGGPPGAEPQ 1153

RESULT 5
US-08-605-672-3
Sequence 3, Application US/08605672
Patent No. 5817515
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. 5817515el Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/605,672
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994

ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32684
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1153 amino acids
TYPE: amino acid
STRADEDNESS: single
TOPOLOGY: linear
MOLSCULE TYPE: protein
US-08-605-672-3
Query Match 100.0%; Score 5956; DB 2; Length 1153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALRVLLTALTLCHGNLDTENAMTFQENARGFCQSVVQLQSGRVVVGAPQEIIVAANOR 60
DB 1 MALRVLLTALTLCHGNLDTENAMTFQENARGFCQSVVQLQSGRVVVGAPQEIIVAANOR 60
QY 61 GSLYQCDYSTGSCCEPIRLQVPEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSNTYVK 120
DB 61 GSLYQCDYSTGSCCEPIRLQVPEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSNTYVK 120
QY 121 GLCFPLGSLNLRQOPKQPEALRGCCQEDSDIAFLIDGSGSIIPHDFRMRKEFVSTVMEQL 180
DB 121 GLCFPLGSLNLRQOPKQPEALRGCCQEDSDIAFLIDGSGSIIPHDFRMRKEFVSTVMEQL 180
QY 181 KSKTFLSLMQSBEFRIHFTFKPEQNNPNPSLVKPKITQLGRTHATATGIRKVVRELFN 240
DB 181 KSKTFLSLMQSBEFRIHFTFKPEQNNPNPSLVKPKITQLGRTHATATGIRKVVRELFN 240
QY 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVITPEADREGVIRVYGVGDAPFSEKSRQEL 300
DB 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVITPEADREGVIRVYGVGDAPFSEKSRQEL 300
QY 301 NTIASKPPRDHVFNQNNPEALKTIQNLREKIFALTEGTQTGSSSSPHEHMSQEGFSAAT 360
DB 301 NTIASKPPRDHVFNQNNPEALKTIQNLREKIFALTEGTQTGSSSSPHEHMSQEGFSAAT 360
QY 361 SNGPLLSTVGSYDMAGGVFLYTSKSKSTFINNTRVDSMDNDAYLGYAAAIILNRVQSLV 420
DB 361 SNGPLLSTVGSYDMAGGVFLYTSKSKSTFINNTRVDSMDNDAYLGYAAAIILNRVQSLV 420
QY 421 LGAPRYQHIGLVAMPFRONTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAP 480
DB 421 LGAPRYQHIGLVAMPFRONTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAP 480
QY 481 HYYEQTRGGQSVCPPLRGQARWQCDVLYGEOQPMGRPGAALTVLGDVNGDKLTDVA 540
DB 481 HYYEQTRGGQSVCPPLRGQARWQCDVLYGEOQPMGRPGAALTVLGDVNGDKLTDVA 540
QY 541 ICAPGEONRGAVLPHGTGSGSISPHSORSITAGSKLSPLOYFGQSLGGQDLTMDGLV 600
DB 541 ICAPGEONRGAVLPHGTGSGSISPHSORSITAGSKLSPLOYFGQSLGGQDLTMDGLV 600
QY 601 DLTVGAQGHVLLRSQPVLRVKAIMEFNPRAVARNVFCNDQVVKGEAGEVRVCLHVOK 660
DB 601 DLTVGAQGHVLLRSQPVLRVKAIMEFNPRAVARNVFCNDQVVKGEAGEVRVCLHVOK 660
QY 661 STRDLREGQIQSVVYDIALDSGRPHSRAVFNETKSTRQVGLTQTCTETLKLQLP 720
DB 661 STRDLREGQIQSVVYDIALDSGRPHSRAVFNETKSTRQVGLTQTCTETLKLQLP 720
QY 721 NCIEDPVSPIVLRNLSVGTPLSAFGNLRPVLAEDAQRULFTALPFPEKNCNDNICQDD 780
DB 721 NCIEDPVSPIVLRNLSVGTPLSAFGNLRPVLAEDAQRULFTALPFPEKNCNDNICQDD 780

Qy 781 LSITFSFMSLDCVVGPPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQORS 840
Db 781 LSITFSFMSLDCVVGPPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQORS 840
Qy 841 QRSWELACESASTVSGALKSTSCINHPFPENSEVTFNITFDVDSKASLGNKLLKA 900
Db 841 QRSWELACESASTVSGALKSTSCINHPFPENSEVTFNITFDVDSKASLGNKLLKA 900
Qy 901 NVTSENNPRTNKTFFQLELPVKYAVYVMTVSHGVSTKYLNTFTASENTRVYMQHGYOVSN 960
Db 901 NVTSENNPRTNKTFFQLELPVKYAVYVMTVSHGVSTKYLNTFTASENTRVYMQHGYOVSN 960
Qy 961 LQORSLPISLVLVRLNQVIMDRPQVTSSENLSSTCHTKERLPDSHDFLAEIRKAPV 1020
Db 961 LQORSLPISLVLVRLNQVIMDRPQVTSSENLSSTCHTKERLPDSHDFLAEIRKAPV 1020
Qy 1021 VNCSTAVCQRIQCDIPFGIOBEFNATLKGNLSPDWYIKTSHNHLIIIVSTAIIIFNDSVF 1080
Db 1021 VNCSTAVCQRIQCDIPFGIOBEFNATLKGNLSPDWYIKTSHNHLIIIVSTAIIIFNDSVF 1080
Qy 1081 TLLPQGGAFVRSQTSKVEPEVFNPLPLIVGSSVGGLLLLALITAAALYKLGFFKRYXD 1140
Db 1081 TLLPQGGAFVRSQTSKVEPEVFNPLPLIVGSSVGGLLLLALITAAALYKLGFFKRYXD 1140
Qy 1141 MMSEGGPPGABQ 1153
Db 1141 MMSEGGPPGABQ 1153

RESULT 6
US-08-482-293A-3
Sequence 3, Application US/08482293A
Patent No. 5831029
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. 5831029e1 Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESSES: Marehall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482,293A
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32684
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:

LENGTH: 1153 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-482-293A-3
Query Match
Best Local Similarity 100.0%; Score 5956; DB 2; Length 1153;
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MALRVLLLTALTICHTGNLDENAMTQENARFGQSVVQLQGRVGVVGAPOQIVANQR 60
Db 1 MALRVLLLTALTICHTGNLDENAMTQENARFGQSVVQLQGRVGVVGAPOQIVANQR 60
Qy 61 GSIYQCDYSGSCPTLRLOVPEAVNMSGLSLAATTPPOLACGPTVHQTCEHTYVK 120
Db 61 GSIYQCDYSGSCPTLRLOVPEAVNMSGLSLAATTPPOLACGPTVHQTCEHTYVK 120
Qy 121 GLCFLFGSNLRQOPQKPEALRGCPQEDSDIAFLIDSGSIIIPDFRMEKPFYSTVMEQL 180
Db 121 GLCFLFGSNLRQOPQKPEALRGCPQEDSDIAFLIDSGSIIIPDFRMEKPFYSTVMEQL 180
Qy 181 KXSKTLFSLMOYSEEFRIHFTKPEFQNNPNSRLVKPIITOLLGRTHATGIRKVVRELFN 240
Db 181 KXSKTLFSLMOYSEEFRIHFTKPEFQNNPNSRLVKPIITOLLGRTHATGIRKVVRELFN 240
Qy 241 ITNGARKNAFKILVITDGEKPGDPLGYEDVIPEADREGVIRYVIGVDAFRSEKSRQEL 300
Db 241 ITNGARKNAFKILVITDGEKPGDPLGYEDVIPEADREGVIRYVIGVDAFRSEKSRQEL 300
Qy 301 NTIASKPPRDHVPQNNFALKTIQNLREKIPIAIEGTOTGSSSSSEHEMSQSGFSAIT 360
Db 301 NTIASKPPRDHVPQNNFALKTIQNLREKIPIAIEGTOTGSSSSSEHEMSQSGFSAIT 360
Qy 361 SNGPLLSITVGSYDWAGVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAIILNRVQSLV 420
Db 361 SNGPLLSITVGSYDWAGVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAIILNRVQSLV 420
Qy 421 LGAPRYQHIGLVAMPFRONTGMESNANVKTQIGAYFGASLCSDVDNSGSTDVLIGAP 480
Db 421 LGAPRYQHIGLVAMPFRONTGMESNANVKTQIGAYFGASLCSDVDNSGSTDVLIGAP 480
Qy 481 HYTEQTRGGQVSVCPILPRGQARQWCDVLYGQGGQPGWGFGAALTIVLGVNDGKLTDA 540
Db 481 HYTEQTRGGQVSVCPILPRGQARQWCDVLYGQGGQPGWGFGAALTIVLGVNDGKLTDA 540
Qy 541 IGAPGEEDNRGAVYLFHGTSGSGISPSHSORIASKLSPLQYFGQSLSGQDLTMDGLV 600
Db 541 IGAPGEEDNRGAVYLFHGTSGSGISPSHSORIASKLSPLQYFGQSLSGQDLTMDGLV 600
Qy 601 DLTWGAQGHVLLLRISQPLVRVKAIMEFNPREVARNVPECNDQVVKGEAGEVRVCLHVQK 660
Db 601 DLTWGAQGHVLLLRISQPLVRVKAIMEFNPREVARNVPECNDQVVKGEAGEVRVCLHVQK 660
Qy 661 STDRRLREGQIQSVTVYDLALDSGRPHSRVAVFNETKNSRTRQVTLGLTQTCETLKLQLP 720
Db 661 STDRRLREGQIQSVTVYDLALDSGRPHSRVAVFNETKNSRTRQVTLGLTQTCETLKLQLP 720
Qy 721 NCIEDPVSPIVLRNLSVCTPLSAFNLPEVLAEDAQRLEFALPPEKKGCHNDICQDD 780
Db 721 NCIEDPVSPIVLRNLSVCTPLSAFNLPEVLAEDAQRLEFALPPEKKGCHNDICQDD 780
Qy 781 LSITFSFMSLDCVVGPPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQORS 840
Db 781 LSITFSFMSLDCVVGPPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQORS 840
Qy 841 QRSWELACESASTVSGALKSTSCINHPFPENSEVTFNITFDVDSKASLGNKLLKA 900
Db 841 QRSWELACESASTVSGALKSTSCINHPFPENSEVTFNITFDVDSKASLGNKLLKA 900
Qy 901 NVTSENNPRTNKTFFQLELPVKYAVYVMTVSHGVSTKYLNTFTASENTRVYMQHGYOVSN 960

Db 901 NTSNNPRNKTQFQLELFPVKAVYVWVTSBGVSTKYLNFTASNTSRVQHOYQVSN 960
 Qy 961 LGQSLPISLVFLVRLNQVINDRPQVTFSENLSSTCHTKERLPKSHSDFLAELRKAPV 1020
 Db 961 LGQSLPISLVFLVRLNQVINDRPQVTFSENLSSTCHTKERLPKSHSDFLAELRKAPV 1020
 Qy 1021 VNCIAVCORTQCDIPFPGIQQEENATLKGMLSDWYIKTSHNHLIVSTAEILFNDVSF 1080
 Db 1021 VNCIAVCORTQCDIPFPGIQQEENATLKGMLSDWYIKTSHNHLIVSTAEILFNDVSF 1080
 Qy 1081 TLLPQCGAFVRSQVETKVEPEVENPLPLIVGSSVGGLLILALITAAALYKLGPFKQVYKD 1140
 Db 1081 TLLPQCGAFVRSQVETKVEPEVENPLPLIVGSSVGGLLILALITAAALYKLGPFKQVYKD 1140
 Qy 1141 MMSEGGPPGASPO 1153
 Db 1141 MMSEGGPPGASPO 1153

RESULT 7

US-08-943-363-3
 ; Sequence 3, Application US/08943363
 ; Patent No. 5837478
 ; GENERAL INFORMATION:
 ; APPLICANT: Gallatin, W. Michael
 ; TITLE OF INVENTION: No. 5837478el Human 2 Integrin Alpha Subunit
 ; NUMBER OF SEQUENCES: 114
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
 ; STREET: 233 South Wacker Drive, 6300 Sear Tower
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: United States
 ; ZIP: 60606-6402
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/943,363
 ; FILING DATE:
 ; CLASSIFICATION: 530
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/173,497
 ; FILING DATE: 23-DEC-1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/286,889
 ; FILING DATE: 5-AUG-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/362,652
 ; FILING DATE: 21-DEC-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Williams Jr., Joseph A.
 ; REGISTRATION NUMBER: 38,659
 ; REFERENCE/DOCKET NUMBER: 27866/32684
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 312-474-6300
 ; TELEFAX: 312-474-0448
 ; TELEX: 25-3856
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1153 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-943-363-3

Query Match 100.0%; Score 5956; DB 2; Length 1153;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MALRVLLLTALTLCGPNLDTENAMTFQENARGFGQSVVQLQSGSRVVVGAPQBIVAANOR 60
 Db 1 MALRVLLLTALTLCGPNLDTENAMTFQENARGFGQSVVQLQSGSRVVVGAPQBIVAANOR 60
 Qy 61 GSIYQCDYSGSCERPLRLOUPVEAVNSGLSLAATTSPPQILLACGTPVHQTCSNTYVK 120
 Db 61 GSIYQCDYSGSCERPLRLOUPVEAVNSGLSLAATTSPPQILLACGTPVHQTCSNTYVK 120
 Qy 121 GLCFLFGSNLRQOQKPFPEALRCQPOEDSDIAFLIDSGSGSIIIPHDFRMKEFVSTVMEQL 180
 Db 121 GLCFLFGSNLRQOQKPFPEALRCQPOEDSDIAFLIDSGSGSIIIPHDFRMKEFVSTVMEQL 180
 Qy 181 KKSKTILFSLMOYSEEFRIHETFKFQNNPNRSLVKPITOLLGRTHTATGIRKVRLEFN 240
 Db 181 KKSKTILFSLMOYSEEFRIHETFKFQNNPNRSLVKPITOLLGRTHTATGIRKVRLEFN 240
 Qy 241 ITNGARKNAFKILVWITDGEKFGDPLGYEDVIFEADREGVIRVIVGDAFRSEKSRQEL 300
 Db 241 ITNGARKNAFKILVWITDGEKFGDPLGYEDVIFEADREGVIRVIVGDAFRSEKSRQEL 300
 Qy 301 NTIASKPRDHVQVNNPFAKTIQNLREKIPAIETGTQSGSSSFEHEMSQGFSAAIT 360
 Db 301 NTIASKPRDHVQVNNPFAKTIQNLREKIPAIETGTQSGSSSFEHEMSQGFSAAIT 360
 Qy 361 SNGPILLSVGSYDWAGGVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLV 420
 Db 361 SNGPILLSVGSYDWAGGVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLV 420
 Qy 421 LGAPRYCHIGLVAMFRONTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAP 480
 Db 421 LGAPRYCHIGLVAMFRONTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAP 480
 Qy 481 HYETQTRGGQVSVCPPLRGORARWQCDAYLYGQOGQWGRFGAALTIVLGVNCGDKLTDVA 540
 Db 481 HYETQTRGGQVSVCPPLRGORARWQCDAYLYGQOGQWGRFGAALTIVLGVNCGDKLTDVA 540
 Qy 541 IGAPGEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSCQDQTLMDGLV 600
 Db 541 IGAPGEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSCQDQTLMDGLV 600
 Qy 601 DLTVGAGCHVLLARSQPLRVKAIMFNPREVARNVPECNDQVVKGEAGEVRVCLHVQK 660
 Db 601 DLTVGAGCHVLLARSQPLRVKAIMFNPREVARNVPECNDQVVKGEAGEVRVCLHVQK 660
 Qy 661 STRDLRREGQIQSVVTVYDLALDSGRPHSAVFNKSTRRTQVGLGLTQTCETLKLQLP 720
 Db 661 STRDLRREGQIQSVVTVYDLALDSGRPHSAVFNKSTRRTQVGLGLTQTCETLKLQLP 720
 Qy 721 NCIEDPVSPIVLRNLFSLVGTPLSAFNLAPVLAEDAQRPLTALFPPEKNCNDNLCQDD 780
 Db 721 NCIEDPVSPIVLRNLFSLVGTPLSAFNLAPVLAEDAQRPLTALFPPEKNCNDNLCQDD 780
 Qy 781 LSITTFMSLDCLVVGPRFNVTVVRNDGDSYRTQVTFPPFDLSYRKVSTLQORS 840
 Db 781 LSITTFMSLDCLVVGPRFNVTVVRNDGDSYRTQVTFPPFDLSYRKVSTLQORS 840
 Qy 841 QRSNRLACESASSTEVSGALKSTSCSINHPIIPENSEVTNITFDVDSKASLGNKLLKA 900
 Db 841 QRSNRLACESASSTEVSGALKSTSCSINHPIIPENSEVTNITFDVDSKASLGNKLLKA 900
 Qy 901 NVTSENNMPRTNTEFQLELPVKYAVVWVTSBGVSTKYLNFTASNTSRVQHOYQVSN 960
 Db 901 NVTSENNMPRTNTEFQLELPVKYAVVWVTSBGVSTKYLNFTASNTSRVQHOYQVSN 960
 Qy 961 LGQSLPISLVFLVRLNQVINDRPQVTFSENLSSTCHTKERLPKSHSDFLAELRKAPV 1020
 Db 961 LGQSLPISLVFLVRLNQVINDRPQVTFSENLSSTCHTKERLPKSHSDFLAELRKAPV 1020
 Qy 1021 VNCIAVCORTQCDIPFPGIQQEENATLKGMLSDWYIKTSHNHLIVSTAEILFNDVSF 1080
 Db 1021 VNCIAVCORTQCDIPFPGIQQEENATLKGMLSDWYIKTSHNHLIVSTAEILFNDVSF 1080

```

QY 1081 TLLPGCGAFVRSQETKVPPEVNPPLPLIVGSSVGGLLLLALITAAALYKLGFFFKROYKD 1140
DB 1081 TLLPGCGAFVRSQETKVPPEVNPPLPLIVGSSVGGLLLLALITAAALYKLGFFFKROYKD 1140
QY 1141 MMSEGGPPGAPQ 1153
DB 1141 MMSEGGPPGAPQ 1153

RESULT 8
US-09-193-043-3
; Sequence 3, Application US/09193043
; Patent No. 6251395
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; TITLE OF INVENTION: No. 6251395el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/193,043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-193-043-3

Query Match 100.0%; Score 5956; DB 3; Length 1153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAURLVLLTALTICGPNLDTENAMTFQENARFGQSVVQLOGSRVVVVGAPQEIIVANOR 60
DB 1 MAURLVLLTALTICGPNLDTENAMTFQENARFGQSVVQLOGSRVVVVGAPQEIIVANOR 60
QY 61 GSIYQCDYSTGSCPIRLQVPEAVNMSLGLSLAATSPQLLACGPTVHQTCSENTYVK 120
DB 61 GSIYQCDYSTGSCPIRLQVPEAVNMSLGLSLAATSPQLLACGPTVHQTCSENTYVK 120
QY 121 GLCFLGSLNRQOPQKPEALRCQPEDSDIAFLIDSGSIIIPHDPRMKPEFVSTWEOQL 180
DB 121 GLCFLGSLNRQOPQKPEALRCQPEDSDIAFLIDSGSIIIPHDPRMKPEFVSTWEOQL 180
QY 181 KXSKTLFSLMOYSEEPRIHPTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFN 240
DB 181 KXSKTLFSLMOYSEEPRIHPTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFN 240
QY 241 ITWGARKNAKILWITDGEKFGDPLGYEDVIEADREGVIRVIGVDAFRSEKROEL 300
DB 241 ITWGARKNAKILWITDGEKFGDPLGYEDVIEADREGVIRVIGVDAFRSEKROEL 300
QY 301 NTIASKPPRDHVQVNNFEALKTIQNLREKIFAIEGTQTGSSSSPEHEMSQSGFSAIT 360
DB 301 NTIASKPPRDHVQVNNFEALKTIQNLREKIFAIEGTQTGSSSSPEHEMSQSGFSAIT 360
QY 361 SNGPELLSTVGSYDWAGGVFLYTSKSKSTFNNTRVDSMDNDVILGYAAAILIENRVQSLV 420
DB 361 SNGPELLSTVGSYDWAGGVFLYTSKSKSTFNNTRVDSMDNDVILGYAAAILIENRVQSLV 420
QY 421 LGAPRYCHIGLAVFRONTGWMESNANVKQTQICAYFGASLCSDVDVDSNGSTDVILIGAP 480
DB 421 LGAPRYCHIGLAVFRONTGWMESNANVKQTQICAYFGASLCSDVDVDSNGSTDVILIGAP 480
QY 481 HYEBQTRGGQVSVCPPLPRQARWQCDVAVLYGCGQFWGRFGAALTIVLGDVNGDKLTDVA 540

```

RESULT 9

```

US-09-688-307A-3
; Sequence 3, Application US/09688307A
; Patent No. 6432404
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 6432404el Human Beta-2
; FILE REFERENCE: 27866/36646
; CURRENT APPLICATION NUMBER: US/09/688,307A
; CURRENT FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/605,672
; PRIOR FILING DATE: 1996-02-22
; PRIOR APPLICATION NUMBER: 08/173,497
; PRIOR FILING DATE: 1993-12-23
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent In Ver. 2.0

```

```

DB 481 HYEBQTRGGQVSVCPPLPRQARWQCDVAVLYGCGQFWGRFGAALTIVLGDVNGDKLTDVA 540
QY 541 IGAPGEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSRLOFTQGSLSGGQDITMDGLV 600
DB 541 IGAPGEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSRLOFTQGSLSGGQDITMDGLV 600
QY 601 DLTVGAQGHVLLRSQPVLRUKAIMEFNPREVARNVFECDNDVVKGEAGVRVCLHVOK 660
DB 601 DLTVGAQGHVLLRSQPVLRUKAIMEFNPREVARNVFECDNDVVKGEAGVRVCLHVOK 660
QY 661 STRDLREGQIQSVVTVYDLALDSCRPHSRVAVNFETKSTRQOTQVLGLTQTCETLKLQLP 720
DB 661 STRDLREGQIQSVVTVYDLALDSCRPHSRVAVNFETKSTRQOTQVLGLTQTCETLKLQLP 720
QY 721 NCIEDPVSPIVLRNPLNSLVGTPLSAPGNLRDVAEDAQRLFTALFPPEKNCNDNICODD 780
DB 721 NCIEDPVSPIVLRNPLNSLVGTPLSAPGNLRDVAEDAQRLFTALFPPEKNCNDNICODD 780
QY 781 LSITFSFMSLDCLVVGSPREENVTVTVRNDCGDSYRTQVTFPPFLDLISYRKVSTLQORS 840
DB 781 LSITFSFMSLDCLVVGSPREENVTVTVRNDCGDSYRTQVTFPPFLDLISYRKVSTLQORS 840
QY 841 QRSWLACESASSTEVSGALKSTSCSINHPIPPENSEVTENITFDVDSKASLGNKLLKA 900
DB 841 QRSWLACESASSTEVSGALKSTSCSINHPIPPENSEVTENITFDVDSKASLGNKLLKA 900
QY 901 NVTSENMPRTNKTEFQLELPVKYAVVMVTVSHGVSTKYLNFATASENTSRVMOHQYOVSN 960
DB 901 NVTSENMPRTNKTEFQLELPVKYAVVMVTVSHGVSTKYLNFATASENTSRVMOHQYOVSN 960
QY 961 LGQRLSLPISLVPLVPLRNQTVIWDRCQVTSSEMLSSSTCHTKERLPSHSDFLAELRKAPV 1020
DB 961 LGQRLSLPISLVPLVPLRNQTVIWDRCQVTSSEMLSSSTCHTKERLPSHSDFLAELRKAPV 1020
QY 1021 VNCSTAVCORIQCDIPPFQIOEFNATLKGSLFDWYIKTSHNLLIIVSTAELIFNDSVF 1080
DB 1021 VNCSTAVCORIQCDIPPFQIOEFNATLKGSLFDWYIKTSHNLLIIVSTAELIFNDSVF 1080
QY 1081 TLLPQCGAFVRSQETKVPPEVNPPLPLIVGSSVGGLLLLALITAAALYKLGFFFKROYKD 1140
DB 1081 TLLPQCGAFVRSQETKVPPEVNPPLPLIVGSSVGGLLLLALITAAALYKLGFFFKROYKD 1140
QY 1141 MMSEGGPPGAPQ 1153
DB 1141 MMSEGGPPGAPQ 1153

```

; SEQ ID NO 3									
; LENGTH: 1153									
; TYPE: PRT									
; ORGANISM: Homo sapiens									
US-09-688-307A-3									
Query Match 100.0%; Score 5956; DB 4; Length 1153;									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	MALRVLLTALTLC	CHGNLDTENAMTFQENARGFGQSVVQLOGSRVVVGAPQEIIVANOR	60					
Db	1	MALRVLLTALTLC	CHGNLDTENAMTFQENARGFGQSVVQLOGSRVVVGAPQEIIVANOR	60					
Qy	61	GSLYQCDYSG	CEPIRLQVPEAVNMSLGLSLAATTSPPQLLACGPTVHQTCE	120					
Db	61	GSLYQCDYSG	CEPIRLQVPEAVNMSLGLSLAATTSPPQLLACGPTVHQTCE	120					
Qy	121	GLCFLFGSNLR	QOQPFPEALRGCPQEDSDIAFLIDSGSIIIPHDFRMKEFVSTVMEQL	180					
Db	121	GLCFLFGSNLR	QOQPFPEALRGCPQEDSDIAFLIDSGSIIIPHDFRMKEFVSTVMEQL	180					
Qy	181	KSKTFLSLMOY	SEEFRIHFTTFKEFQNNPNRSLVKPIITQLLGRTHATGIRKVVRELFN	240					
Db	181	KSKTFLSLMOY	SEEFRIHFTTFKEFQNNPNRSLVKPIITQLLGRTHATGIRKVVRELFN	240					
Qy	241	ITNGARKNAFK	ILVITDGEKFGDPLGYEDVPEADREGVIRVYVIGVGDAPRSEKSRQEL	300					
Db	241	ITNGARKNAFK	ILVITDGEKFGDPLGYEDVPEADREGVIRVYVIGVGDAPRSEKSRQEL	300					
Qy	301	NTIASKPRDHV	QVNNFEALKTIONLREKIPAIETGTQTGSSSSPHEMSQSGFSAIT	360					
Db	301	NTIASKPRDHV	QVNNFEALKTIONLREKIPAIETGTQTGSSSSPHEMSQSGFSAIT	360					
RESULT 10									
US-09-350-259-3									
; Sequence 3, Application US/09350259									
; Patent No. 6620915									
; GENERAL INFORMATION:									
; APPLICANT: Gallatin, Michael W.									
; APPLICANT: Van der Vieren, Monica									
; TITLE OF INVENTION: No. 6620915el Human 2									
; FILE REFERENCE: 27866/35004									
; CURRENT APPLICATION NUMBER: US/09/350,259									
; CURRENT FILING DATE: 1999-07-08									
; EARLIER APPLICATION NUMBER: 09/193,043									
; EARLIER FILING DATE: 1998-11-16									
; EARLIER APPLICATION NUMBER: 08/173,497									
; EARLIER FILING DATE: 1993-12-23									
; EARLIER APPLICATION NUMBER: 08/286,889									
; EARLIER FILING DATE: 1994-08-05									
; EARLIER APPLICATION NUMBER: 08/362,652									
; EARLIER FILING DATE: 1994-12-21									
; EARLIER APPLICATION NUMBER: 08/943,363									
; EARLIER FILING DATE: 1997-10-03									
; NUMBER OF SEQ ID NOS: 114									
; SOFTWARE: Patentin Ver. 2.0									
; SEQ ID NO 3									
; LENGTH: 1153									
; TYPE: PRT									
; ORGANISM: Homo sapiens									
US-09-350-259-3									
Query Match 100.0%; Score 5956; DB 4; Length 1153;									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	MALRVLLTALTLC	CHGNLDTENAMTFQENARGFGQSVVQLOGSRVVVGAPQEIIVANOR	60					
Db	1	MALRVLLTALTLC	CHGNLDTENAMTFQENARGFGQSVVQLOGSRVVVGAPQEIIVANOR	60					
Qy	61	GSLYQCDYSG	CEPIRLQVPEAVNMSLGLSLAATTSPPQLLACGPTVHQTCE	120					
Db	61	GSLYQCDYSG	CEPIRLQVPEAVNMSLGLSLAATTSPPQLLACGPTVHQTCE	120					
Qy	121	GLCFLFGSNLR	QOQPFPEALRGCPQEDSDIAFLIDSGSIIIPHDFRMKEFVSTVMEQL	180					
Db	121	GLCFLFGSNLR	QOQPFPEALRGCPQEDSDIAFLIDSGSIIIPHDFRMKEFVSTVMEQL	180					
Qy	181	KSKTFLSLMOY	SEEFRIHFTTFKEFQNNPNRSLVKPIITQLLGRTHATGIRKVVRELFN	240					
Db	181	KSKTFLSLMOY	SEEFRIHFTTFKEFQNNPNRSLVKPIITQLLGRTHATGIRKVVRELFN	240					
Qy	241	ITNGARKNAFK	ILVITDGEKFGDPLGYEDVPEADREGVIRVYVIGVGDAPRSEKSRQEL	300					
Db	241	ITNGARKNAFK	ILVITDGEKFGDPLGYEDVPEADREGVIRVYVIGVGDAPRSEKSRQEL	300					
Qy	301	NTIASKPRDHV	QVNNFEALKTIONLREKIPAIETGTQTGSSSSPHEMSQSGFSAIT	360					
Db	301	NTIASKPRDHV	QVNNFEALKTIONLREKIPAIETGTQTGSSSSPHEMSQSGFSAIT	360					
Qy	361	SNGLLSTVGS	YDWAGGVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAILLRNVQSLV	420					
Db	361	SNGLLSTVGS	YDWAGGVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAILLRNVQSLV	420					
Qy	421	IGAPRYOHGL	VAFRONTGWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLGAP	480					
Db	421	IGAPRYOHGL	VAFRONTGWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLGAP	480					
Qy	481	HYETQTRGG	QSVVCLPRGQARQCDVLYGEQGPWGRFGAALTIVLGDVNGDKLTDVA	540					
Db	481	HYETQTRGG	QSVVCLPRGQARQCDVLYGEQGPWGRFGAALTIVLGDVNGDKLTDVA	540					
Qy	541	IGAPGEDNR	GAVALFHGTSGSGISPSHSQRIAGSKLSPRLQVFGQSLSGQDLTMDGLV	600					
Db	541	IGAPGEDNR	GAVALFHGTSGSGISPSHSQRIAGSKLSPRLQVFGQSLSGQDLTMDGLV	600					
Qy	601	DLTVGAQGH	VLLRSQVLRVKAIMEFNPREVARNFECNDQVVKGEAGEVRVCLHVQK	660					
Db	601	DLTVGAQGH	VLLRSQVLRVKAIMEFNPREVARNFECNDQVVKGEAGEVRVCLHVQK	660					
Qy	661	STRDLREGQ	TQSVVTVYDLALDSGRPHSRVFNFTKSTRQTOVLGLTQTCETLKLQLP	720					
Db	661	STRDLREGQ	TQSVVTVYDLALDSGRPHSRVFNFTKSTRQTOVLGLTQTCETLKLQLP	720					
Qy	721	NCIEDPVS	PVLRNFSLVGTPLSAFGLNLRVLAEDAQRUFTALFPFKKNGNDNICDD	780					
Db	721	NCIEDPVS	PVLRNFSLVGTPLSAFGLNLRVLAEDAQRUFTALFPFKKNGNDNICDD	780					
Qy	781	LSITFSFMS	LDCLVVGGPREFNVTVVRNDEDSYRTQVTFPPPLDLSYRKVSTLQNR	840					
Db	781	LSITFSFMS	LDCLVVGGPREFNVTVVRNDEDSYRTQVTFPPPLDLSYRKVSTLQNR	840					
Qy	841	QRSWLACES	ASSSTEVSGALAKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLILKA	900					
Db	841	QRSWLACES	ASSSTEVSGALAKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLILKA	900					
Qy	901	NVTSENMPR	NTKTEFQLELPVKYAVVMVTVSHGVSTKYLNFMTASENTSRVMQHQYQVSN	960					
Db	901	NVTSENMPR	NTKTEFQLELPVKYAVVMVTVSHGVSTKYLNFMTASENTSRVMQHQYQVSN	960					

Db 301 NTIASKPPDRHVQVNNFEALKTIONQLREKIFAIEGTQTGSSSSPEHEMSQEGFSAIT 360
Qy 361 SNGPILLSTVGSYDWAAGGVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAIILNRVQSLV 420
Db 361 SNGPILLSTVGSYDWAAGGVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAIILNRVQSLV 420
Qy 421 LGAPRYQHIGLVAMFRQNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAP 480
Db 421 LGAPRYQHIGLVAMFRQNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAP 480
Qy 481 HYEQTRGGQSVSCPLPRGQARWQCDVAVLYGEOQPMGRFGAALTVLGDVNGDKLTDVA 540
Db 481 HYEQTRGGQSVSCPLPRGQARWQCDVAVLYGEOQPMGRFGAALTVLGDVNGDKLTDVA 540
Qy 541 IGAPEDENRGAIVLPHGTSGSGTSPSHSQRIAGSKLSPRLOYGOSLGGQDITMDGLV 600
Db 541 IGAPEDENRGAIVLPHGTSGSGTSPSHSQRIAGSKLSPRLOYGOSLGGQDITMDGLV 600
Qy 601 DLTVAQGHVLLRSQVLRVKAIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQK 660
Db 601 DLTVAQGHVLLRSQVLRVKAIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQK 660
Qy 661 STRDLREGQIQSVVYDIALDSGRPHSRVFNETHKSTRTOVLGLTQTCETLKLQLP 720
Db 661 STRDLREGQIQSVVYDIALDSGRPHSRVFNETHKSTRTOVLGLTQTCETLKLQLP 720
Qy 721 NCIEDPVSPIVLRNLSVGTPLSAFGNLRVLAEDAQRLEFTALFPPEKNCNDNIQDD 780
Db 721 NCIEDPVSPIVLRNLSVGTPLSAFGNLRVLAEDAQRLEFTALFPPEKNCNDNIQDD 780
Qy 781 LSITFSMSLCLVVGGRPRENVTVVNDGEDSYRTQVTFPPFLDLSYRKVSTLQNR 840
Db 781 LSITFSMSLCLVVGGRPRENVTVVNDGEDSYRTQVTFPPFLDLSYRKVSTLQNR 840
Qy 841 QRSRLACESASTSVSALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKA 900
Db 841 QRSRLACESASTSVSALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKA 900
Qy 901 NVTSENMPRTKTEFOLELPVKAVVMTVSHGVSTKYLNTASENTRVMAHQYQVSN 960
Db 901 NVTSENMPRTKTEFOLELPVKAVVMTVSHGVSTKYLNTASENTRVMAHQYQVSN 960
Qy 961 LGQSLPLSLVFLVPLVRLNQVIVDRPOVTFSENLSSTCHTKERLPSHSDPLAELRKPV 1020
Db 961 LGQSLPLSLVFLVPLVRLNQVIVDRPOVTFSENLSSTCHTKERLPSHSDPLAELRKPV 1020
Qy 1021 VNCISAVCQRIQCDIPFPGIQEEFNATLKGNLSDFDWIKTSHNLLIVSTAELFNDSVF 1080
Db 1021 VNCISAVCQRIQCDIPFPGIQEEFNATLKGNLSDFDWIKTSHNLLIVSTAELFNDSVF 1080
Qy 1081 TLLPQGNFVSESOTETKVEPSEVNPPLIVGSSVGGLLLALITAALYKLGPFKQYKD 1140
Db 1081 TLLPQGNFVSESOTETKVEPSEVNPPLIVGSSVGGLLLALITAALYKLGPFKQYKD 1140
Qy 1141 MWSEGGPPGABEQ 1153
Db 1141 MWSEGGPPGABEQ 1153

RESULT 11

US-08-476-062A-43
; Sequence 43, Application US/08476062A
; Patent No. 5877275
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY
; TITLE OF INVENTION: RESPONSES WITH BETA2 INTEGRINS
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; STREET: Fish & Richardson P.C.
; CITY: Boston
; STATE: MA

COUNTRY: US
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,062A
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/216,081
FILING DATE: 21-MAR-1994
APPLICATION NUMBER: 07/637,830
FILING DATE: 04-JAN-1991
APPLICATION NUMBER: 07/539,842
FILING DATE: 18-JUN-1990
APPLICATION NUMBER: 07/212,573
FILING DATE: 28-JUN-1988
ATTORNEY/AGENT INFORMATION:
NAME: Freeman, John W.
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 00786/068003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 1152 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-476-062A-43

Query Match 99.5%; Score 5925.5; DB 2; Length 1152;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1150; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
Qy 1 MALRVLALLTALTLCHGNLDENAMTFQENARGFGQSVVQLQGRVVVGAPQBIIVANQR 60
Db 1 MALRVLALLTALTLCHGNLDENAMTFQENARGFGQSVVQLQGRVVVGAPQBIIVANQR 60
Qy 61 GSLYQCDYTGSCPTIRLQVPEAVNMSGLSLAATTSPLLACQPTVHCTSENYUK 120
Db 61 GSLYQCDYTGSCPTIRLQVPEAVNMSGLSLAATTSPLLACQPTVHCTSENYUK 120
Qy 121 GLCFLPGSNLRQOPKPFPEALRGCPQEDSDIAFLDGSIIIPHDFRRMKFVSTVMEQL 180
Db 121 GLCFLPGSNLRQOPKPFPEALRGCPQEDSDIAFLDGSIIIPHDFRRMKFVSTVMEQL 180
Qy 181 KSKTLLFSLMQYSEBEPRIHFTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFN 240
Db 181 KSKTLLFSLMQYSEBEPRIHFTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFN 240
Qy 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVPEADREGVIRVVGDAFRSEKROEL 300
Db 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVPEADREGVIRVVGDAFRSEKROEL 300
Qy 301 NTIASKPPDRHVQVNNFEALKTIQNQLREKIFAIEGTQTGSSSSPEHEMSQEGFSAIT 360
Db 301 NTIASKPPDRHVQVNNFEALKTIQNQLREKIFAIEGTQTGSSSSPEHEMSQEGFSAIT 360
Qy 361 SNGPILLSTVGSYDWAAGGVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAIILNRVQSLV 420
Db 361 SNGPILLSTVGSYDWAAGGVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAIILNRVQSLV 420
Qy 421 LGAPRYQHIGLVAMFRQNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAP 480
Db 421 LGAPRYQHIGLVAMFRQNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAP 480
Qy 481 HYEQTRGGQSVSCPLPRGQARWQCDVAVLYGEOQPMGRFGAALTVLGDVNGDKLTDVA 540

481 HYTEQTRGGQVSVCLPRG-RARWQCDVILGSGQPGWGFAGALTVLGDVNGKLTJDA 539
541 IGAPGEDNRGAVYLFHGTSGSISHSQRIAGSKLSPALQYFGQSLSGGQSLTWDGLV 600
540 IGAPGEDNRGAVYLFHGTSGSISHSQRIAGSKLSPALQYFGQSLSGGQSLTWDGLV 599
601 DLTVGAQGHVLLRSQVLRVKAIMEFNPVARNVFECDQVVKGEAGEVRVCLHVQK 660
600 DLTVGAQGHVLLRSQVLRVKAIMEFNPVARNVFECDQVVKGEAGEVRVCLHVQK 659
661 STRDLRREGQIQSVVTVYDIALDSGRPHSRVAVNETKNSTRTOTVGLTQTCETLKLQLP 720
660 STRDLRREGQIQSVVTVYDIALDSGRPHSRVAVNETKNSTRTOTVGLTQTCETLKLQLP 719
721 NCTEDPVSPIVLRNLFSLVGTPLSAFNGRLPVLAEQAORLFTALFPFEKXGNDNICODD 780
720 NCTEDPVSPIVLRNLFSLVGTPLSAFNGRLPVLAEQAORLFTALFPFEKXGNDNICODD 779
781 LSITPFSNGLDCLVWGSPREFNVTVRNDGDSYRTQVTFPPPLDLSYRKVSTLQNRK 840
780 LSITPFSNGLDCLVWGSPREFNVTVRNDGDSYRTQVTFPPPLDLSYRKVSTLQNRK 839
841 QRSWLACSSASTSVSGALKSTCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKA 900
840 QRSWLACSSASTSVSGALKSTCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKA 899
901 NVTSENNMPRTNKTPEQLBPKVIAVYVTVTSHGVSTKYNFTASENTSRVQHQVQVSN 960
900 NVTSENNMPRTNKTPEQLBPKVIAVYVTVTSHGVSTKYNFTASENTSRVQHQVQVSN 959
961 LQORSPISLVFLVPLVRLNQTWIDRPQVTFPSENLSSTCHTERLPSSHDFLAEARKAPV 1020
960 LQORSPISLVFLVPLVRLNQTWIDRPQVTFPSENLSSTCHTERLPSSHDFLAEARKAPV 1019
1021 VNCSTAVCQRIQCDIPFFGQIEEFNATLKNLSFDWYIKTSHNHLIVSTABILFNDVSF 1080
1020 VNCSTAVCQRIQCDIPFFGQIEEFNATLKNLSFDWYIKTSHNHLIVSTABILFNDVSF 1079
1081 TLLPGGAFVRSQTEKTFEVPFNPPLPLIVGSSVGLLILLALITAAVYKLGFFPRQYKD 1140
1080 TLLPGGAFVRSQTEKTFEVPFNPPLPLIVGSSVGLLILLALITAAVYKLGFFPRQYKD 1139
1141 MASEGGPPGAEQ 1153
1140 MASEGGPPGAEQ 1152

RESULT 12
PCT-US96-01314-43
Sequence 43, Application PC/TUS9601314
GENERAL INFORMATION:
APPLICANT: M. Amin Arnaout
TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN
TITLE OF INVENTION: ANTAGONISTS
NUMBER OF SEQUENCES: 78
CORRESPONDENCE ADDRESS:
ADDRESSER: Fish & Richardson P.C.
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 50Z or 55SX
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/01314
FILING DATE: 30-JAN-96
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/380,167

FILING DATE: 30-JAN-95
ATTORNEY/AGENT INFORMATION:
NAME: John W. Freeman
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 00786/267001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 1152
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
PCT-US96-01314-43
Query Match 99.5%; Score 5925.5; DB 5; Length 1152;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1150; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
QY 1 MALRVLLLTALTTLCHGFNLDTENAMTFQENARGFQSVVQLQGSRRVVVGAPOEIVAAQR 60
Db 1 MALRVLLLTALTTLCHGFNLDTENAMTFQENARGFQSVVQLQGSRRVVVGAPOEIVAAQR 60
QY 61 GSLYQCDYSTGSCBPIRLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVK 120
Db 61 GSLYQCDYSTGSCBPIRLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVK 120
QY 121 GLCFPLGSLNLQCKQKPEALRGCPQEDSDTAFILDKSGSIIBHDFRMRKPFVSTVMQL 180
Db 121 GLCFPLGSLNLQCKQKPEALRGCPQEDSDTAFILDKSGSIIBHDFRMRKPFVSTVMQL 180
QY 181 KSKTLPFLMYSSEFRIHFTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFN 240
Db 181 KSKTLPFLMYSSEFRIHFTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFN 240
QY 241 ITGARKNAFKILVITDGEKFGDPLGYEDVTPADRGVTRVYVIGCDAPRSEKROEL 300
Db 241 ITGARKNAFKILVITDGEKFGDPLGYEDVTPADRGVTRVYVIGCDAPRSEKROEL 300
QY 301 NTIASKPPDRHVFQVNNFEALKTIONQLREKI PAIEGTQTCSSSPFHEMSEGEFSAAIT 360
Db 301 NTIASKPPDRHVFQVNNFEALKTIONQLREKI PAIEGTQTCSSSPFHEMSEGEFSAAIT 360
QY 361 SNGPLLSTVGSYDWAGGVFLYTSKEKSTFINMTVDSMDNDAYLGYAAAAILRNRVQSLV 420
Db 361 SNGPLLSTVGSYDWAGGVFLYTSKEKSTFINMTVDSMDNDAYLGYAAAAILRNRVQSLV 420
QY 421 LGAPRYQHIGLVAMFRONTGWESNANVKGTQIGAYFGASLCSVDMSNGSTDLVLIGAP 480
Db 421 LGAPRYQHIGLVAMFRONTGWESNANVKGTQIGAYFGASLCSVDMSNGSTDLVLIGAP 480
QY 481 HYTEQTRGGQVSVCLPRGQARWQCDVILGSGQPGWGFAGALTVLGDVNGKLTJDA 540
Db 481 HYTEQTRGGQVSVCLPRG-RARWQCDVILGSGQPGWGFAGALTVLGDVNGKLTJDA 539
QY 541 IGAPCEDNRGAVYLFHGTSGSISHSQRIAGSKLSPALQYFGQSLSGGQSLTWDGLV 600
Db 540 IGAPCEDNRGAVYLFHGTSGSISHSQRIAGSKLSPALQYFGQSLSGGQSLTWDGLV 599
601 DLTVGAQGHVLLRSQVLRVKAIMEFNPVARNVFECDQVVKGEAGEVRVCLHVQK 660
600 DLTVGAQGHVLLRSQVLRVKAIMEFNPVARNVFECDQVVKGEAGEVRVCLHVQK 659
661 STRDLRREGQIQSVVTVYDIALDSGRPHSRVAVNETKNSTRTOTVGLTQTCETLKLQLP 720
660 STRDLRREGQIQSVVTVYDIALDSGRPHSRVAVNETKNSTRTOTVGLTQTCETLKLQLP 719
721 NCTEDPVSPIVLRNLFSLVGTPLSAFNGRLPVLAEQAORLFTALFPFEKXGNDNICODD 780
720 NCTEDPVSPIVLRNLFSLVGTPLSAFNGRLPVLAEQAORLFTALFPFEKXGNDNICODD 779


```

COUNTRY: US
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,622A
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/216,081
FILING DATE: 21-MAR-1994
APPLICATION NUMBER: 07/637,830
FILING DATE: 04-JAN-1991
APPLICATION NUMBER: 07/539,842
FILING DATE: 18-JUN-1990
APPLICATION NUMBER: 07/212,573
FILING DATE: 28-JUN-1988
ATTORNEY/AGENT INFORMATION:
NAME: Freeman, John W.
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 00786/068003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 1163 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-476-062A-44

Query Match 58.8%; Score 3503; DB 2; Length 1163;
Best Local Similarity 61.3%; Pred. No. 3.8e-287;
Matches 700; Conservative 138; Mismatches 297; Indels 6; Gaps 4;

QY 5 VLLTALTLCGHNLDENAMFOENARFGQSVQLOQSRVVGAPQBIIVAAORGSLY 64
DB 8 LLLFTALSLGSLNLTBETLAPRVDSAGFSVQYANVWVGAPQKITAANQTGGLY 67
QY 65 QCDYSTGSCPTLRQVPAVANMSLGLSLAATSPQLLACGPTVHQTCSNTYVKGICF 124
DB 68 QCCYSTGACETICLQVPPAVANMSLGLSLASTTSPSLLACGPTVHHCGRNMYLTGLCF 127
QY 125 LFGSNLRQCPQPEALRCGPQSDSDIAFLIDGSGSIIIPHDFRMKEFPVSTVMBOLKXSK 184
DB 128 LLOPT--QTLPLVSRQCPQEQDQDIVFLIDGSGSISSENATPMQFVRAVISQFRPS 185
QY 185 TLPSLMQYSEEFRIHTPKFQNNPNRSLVXPIITOLLGRHTATGIRKVVRELFNITNG 244
DB 186 TQPSLMQFNSKFKQTHPTPEFRSTNPLSLASVHQLQGFTVTATAIQNVVHRLPHASYG 245
QY 245 ARKNAEKLIVITDGEKFGDPIGYEDVIEADREGVIRYVIGVDAPRSEKSRQELNTIA 304
DB 246 ARDATKILIVITDGEKFGDSDLDYKDVIPNADAAAGIIRYVIGVDAPRSEKSRQELNTIA 305
QY 305 SKPPDRHVQVNNFEALKTIQNLREKIPIAEGTQTSSTSSFEHMSQEGPSAAITNSGP 364
DB 306 SKPSQHEIKFVEDFDALKIQNLKEKIPIAEGTETTSSTSSFEHMSQEGPSAAITNSGP 365
QY 365 LLSVSGSYDWAGGVLYTSEKSTFTNTRVDSMDNDAYLGVAALILNRVQSLVGLAP 424
DB 366 VLGAAGSFTWGGGAFLYPPNMSFTFNNQSNVMDRDSYLGYSTELALWKGVQSLVGLAP 425
QY 425 RYCHIGLVAMFRONTGWMSNANVKGTQICAVFGASLCSVDVDSNGSTDVLIIGAPHYYE 484
DB 426 RYCHIGLVAMFRONTGWMSNANVKGTQICAVFGASLCSVDVDSNGSTDVLIIGAPHYYE 485
QY 485 QTRGGQVSVCPPLRGQRARWQCDVILYGEQGPWGFAGALTIVLGDVNGDKLTDVAIGAP 544
DB 486 QTRGGQVSVCPPLRGQRARWQCDVILYGEQGPWGFAGALTIVLGDVNGDKLTDVAIGAP 544
QY 545 GEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFOOSLSGGODLTMGDLVLTUV 604
DB 545 GEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFOOSLSGGODLTMGDLVLTUV 604
QY 605 GAQGHVLLRSOPVLRYKAIWSENPVARNVPCNDQVVKGEAGEVVRVCLHVQKSTRD 664
DB 605 GARCQVLLLRTRPRVWVGVSQMFIPAIIPRPAFCREQVQVSEQTLVQSNICLYIDKRSKN 664
QY 665 RLREGQIQSVVYDIALDSDGRPHSRVFNENKSTRQQTQVLTGTQTCETLKLQPCNIE 724
DB 665 RLREGQIQSVVYDIALDSDGRPHSRVFNENKSTRQQTQVLTGTQTCETLKLQPCNIE 724
QY 725 DPVSPVILRLNFSVLGTPSLFAGNLRPVLAEDAOELFTALPPEKNCNNDNICODDLSIT 784
DB 725 DSVTPITLRLNFTLVGKREPLAFRLNRLMAMLAQRYFTASLPPEKNCNNDNICODDLSIT 784
QY 785 FSNMLOCLVYGGPRRBNVTVVRNDGEDSYRTQVTFPFLDLSYRKVSTLQNRQSRWS 844
DB 785 FSNMLOCLVYGGPRRBNVTVVRNDGEDSYRTQVTFPFLDLSYRKVSTLQNRQSRWS 844
QY 845 PLACESASSTEVSGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTS 904
DB 845 HLTCDSPVPG--SQTWSTSCRIINELIFRGAQITFLATFDVSPKAVILGDRLLLTANVSS 902
QY 905 ENNMPTNKTEPQLELPVKYAVYVWVTSHGVSSTKYLNPFTAS-ENTSRVMOHQYQVSNLQ 963
DB 903 ENTFRSTKITFQLELPVKYAVYVWVTSHGVSSTKYLNPFTAS-ENTSRVMOHQYQVSNLQ 962
QY 964 RSLPISLVPLVPLNQLQVWDRPOVTFSENLSTCTKRLPSHSDFLAELRKAQVNC 1023
DB 963 RDLFVSINFWVVELNQEAVMVDVESHPPQNPSCSEKSIAPASDFLAHIQKNPVLDC 1022
QY 1024 SIACQRIQCDIPFEGIOEENATLKNLSFDMWIKTSHNLLIVSTAEILFNDSVFTLL 1083
DB 1023 SIACGLRCPDVPSEFVQSELDFTLKNLSFGNWRQILQKVVSVVVAEITFDISVYQL 1082
QY 1084 PQGAFVRSQTEKTVPEPEPNPLPLIVGSSVGGLLLALITALYKUGFKPKQYKQMS 1143
DB 1083 PQGAFVRSQTEKTVPEPEPNPLPLIVGSSVGGLLLALITALYKUGFKPKQYKQMS 1142
QY 1144 B 1144
DB 1143 B 1143

RESULT 15
PCT-US96-01314-44
; Sequence 44, Application PC/TUS9601314
; GENERAL INFORMATION:
; APPLICANT: M. Amin Arnaout
; TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN
; TITLE OF INVENTION: ANTAGONISTS
; NUMBER OF SEQUENCES: 78
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/01314
; FILING DATE: 30-JAN-96
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/380,167
; FILING DATE: 30-JAN-95

```

Search completed: June 7, 2004, 17:19:16
Job time : 23.6066 secs

ATTORNEY/AGENT INFORMATION:		Query Match	
NAME:	John W. Freeman	58.8%;	Score 3503; DB 5; Length 1163;
REGISTRATION NUMBER:	29,066	Best Local Similarity	61.3%; Pred. No. 3,86-287;
REFERENCE/DOCKET NUMBER:	00786/367001	Matches	700; Conservative 138; Mismatches 297; Indels 6; Gaps 4;
TELECOMMUNICATION INFORMATION:			
TELEPHONE:	(617) 542-5070		
TELEFAX:	(617) 542-8906		
TELEX:	200154		
INFORMATION FOR SEQ ID NO: 44:			
SEQUENCE CHARACTERISTICS:			
TYPE:	amino acid		
STRANDEDNESS:	%		
TOPOLOGY:	linear		
PCT-US96-01314-44			
QY	5	VLLLTALTLC	CHGNLDTENAMTPOENARGCQSVVOLQGSRVVVGAPQEIIVANORGSLY 64
DB	8	LLLTALATSL	GNLDTETETAFRVDSAGFGDSVVQYANSVVVGAPQKITAANQTGGLY 67
QY	65	QCDYSTGSCP	IRLOVPEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVKGICF 124
DB	68	QCGYSTGACP	PIGLQVPEAVNMSLGLSLASTTSPSOLLACGPTVHHECGRNWYLTGLCP 127
QY	125	LFGSNLRQQQ	KQKPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQLKSK 184
DB	128	LLGPT--QLT	QLRVSRQECFQBQDIVELIDGSGSISRNFATMNFVRAVISQFORPS 185
QY	185	TLPSLMQVSR	EPRTHFTFKFONPNPNSLVKPTQLLGRTHATGIRKVVURELFNITNG 244
DB	186	TQPSLMQF	SNKFTQHTFTEPRFTSNPLSLASVHQLGFTYTATATQNVVHRLFHASYG 245
QY	245	ARKNAFKILV	ITDGRKPGDPLGYEDVIPLEADREGVIRYIVGVGDAPRSEKSRQELMTIA 304
DB	246	ARRDATKILV	ITDCKEGSDLYKDVIPMAADAAGIRYAIYGVGLAFONRNSMKELANDIA 305
QY	305	SKPRDVRVQV	NFNFEALTIQNLREKIFALEGTQTSSSSFHEMSQEGFSAATITNGP 364
DB	306	SKPSQEHIFK	VEDFDALQIQNLQEKIFALEGTETSSSFELEMQEGFSAVFTDGP 365
QY	365	LLSVGSDYWG	AGGYFLVYTSKEKSTFINNTRVDSQNDAYLGYAAAILNRVQSLVLGAP 424
DB	366	VLGAVGFTWG	SGAFLYPPNMSPTFINNSQENVMDRDSYLGYSTELALWKGVSILVLCAP 425
QY	425	RYQHILGVAM	FRQNTGMESNANYKGTQIAYFGASLCSVDVDVNGSTDLVLI GAPHYYE 484
DB	426	RYQHTGRAVI	QTQVSRQRMKAETVGTQISGYFGASLCSVDVDVNGSTDLVLI GAPHYYE 485
QY	485	QTRGGQVSVCP	LPQRQARWOCDAVLYEQEQQPMGRFGAALTYLGDVNGDKLTDVAIGAP 544
DB	486	QTRGGQVSVCP	LPGRWR--RWWCDAVLYEQGHQHPGRFGAALTDLGVNGDKLTDVVIGAP 544
QY	545	GEEDNRGAVYL	PHGTSGSGGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTV 604
DB	545	GEENRGAVYL	PHGVLGPSISPSHSQRIAGSLSRLQYFGQALSGGQDLTQDGLVDLAV 604
QY	605	GAQGHVLLRS	QPLVRKAIEMFNPREVARNVFCNDQVKGKEAGEVRVCLHVOKSTRD 664
DB	605	GARGQVLLRT	RPVLWGVSWQFIPTABIPRSNFECEQVVSQTLVQSNICLYIDRSKN 664
QY	665	RLREGQIQSV	VTYDIALDSGRPHSAVRNETHKNSTRQTQVLGLTQTCETKLQLPNCIE 724
DB	665	LLGSRDLQSS	VTLDLALDQRLSPRATPQETKNRSLSRVRLGLKAHCENFLLPSCVE 724
QY	725	DPVSDIVRLN	PSTLPSAFGNRPVLAEQAQLFTALPPEFKNCNDNLCODDLISIT 784
DB	725	DSVTITLRLN	PFTLVGKPLAFNRNLRPLMLAAQRYFTASLPPEFKNCNADHITCQNLGIS 784
QY	785	FSFMSLDCLV	VGCGPREFNVTIVTRNDGSDSVRTQVTFPPFDLSYKRVSTIQNORSQS 844

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model
Run on: June 7, 2004, 17:05:16 ; Search time 14.559 Seconds
(without alignments)
7512.163 Million cell updates/sec

Title: US-09-902-481B-3
Perfect score: 5879
Sequence: 1 FNLDTENAMTQENARGPGQ.....FKROYKDMXSEGGPPGAEPQ 1137
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78.*
1: Piri.*
2: Piri2.*
3: Piri3.*
4: Piri4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5839	99.3	1153	1 RWHU1B	cell surface glyco
2	4447	75.6	1153	2 S00551	leukocyte surface
3	3456	58.3	1163	1 RWHU1C	cell surface glyco
4	1532.5	26.1	1170	2 S03308	cell surface glyco
5	1516.5	25.8	1163	2 I56126	lymphocyte function
6	1128	19.2	1179	2 A53213	integrin alpha-E c
7	1085.5	18.5	1151	2 A45226	integrin alpha-1 c
8	1071	18.2	1170	2 I45914	integrin alpha-2 s
9	1057	18.0	1178	2 S44142	VLA-2 protein homo
10	1054	17.9	1181	2 A33998	integrin alpha-2 c
11	1049	17.8	1180	2 A35854	integrin alpha-1 c
12	666	11.3	1039	2 A41131	lymphocyte-Peyer's
13	642	10.9	1038	2 S05046	integrin alpha-4 c
14	630	10.7	1035	2 I58409	integrin alpha-9 c
15	614.5	10.5	1041	2 T31437	integrin alpha cha
16	579.5	9.9	1054	2 JC7294	alphanp integrin -
17	572.5	9.7	1051	2 A35761	cell surface glyco
18	567.5	9.7	1053	2 I55534	VLA-3 alpha subuni
19	555.5	9.4	1053	2 S44250	integrin alpha-5 c
20	551.5	9.4	1034	2 A36108	integrin alpha-v c
21	539	9.2	1044	2 T10050	integrin alpha-6 c
22	533.5	9.1	1072	2 A38457	fibronectin recept
23	532	9.0	1049	2 A27079	integrin alpha-6 c
24	531	9.0	1073	2 B36429	integrin alpha-6 c
25	530	9.0	1048	2 A27421	integrin alpha-5 c
26	529.5	9.0	1051	2 A40021	integrin alpha-3 c
27	524.5	8.9	1091	2 A41543	integrin alpha-6 c
28	514.5	8.8	1044	2 S14516	integrin alpha-8 c
29	497	8.5	1094	2 A29637	position-specific

RESULT 1

RWU1B

cell surface glycoprotein CD11b precursor [validated] - human
N:Alternate names: complement receptor type 3 alpha chain; leukocyte adhesion protein 1
eukocyte integrin alpha chain; neutrophil adherence receptor alphaM chain
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text change 31-Dec-2000
C:Accession: A31108; A28915; A41600; A30892; A32218; A46526; A26091; I52567
R:Corbi, A.L.; Kishimoto, T.K.; Miller, L.J.; Springer, T.A.
J. Biol. Chem. 263, 12403-12411, 1988
A:Title: The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, B.
A:Reference number: A31108; MUID:88315033; PMID:2457584
A:Accession: A31108
A:Molecule type: mRNA
A:Residues: 1-1153 <COR>
A:Cross-references: GB:J03925; NID:G187284; PIDN:AAA59544.1; PID:G307148
A:Note: part of this sequence was confirmed by protein sequencing
J. Cell Biol. 106, 2153-2158, 1988
R:Arnaout, M.A.; Gupta, S.K.; Pierce, M.W.; Tenen, D.G.
A:Title: Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor
A:Reference number: A28915; MUID:88257215; PMID:2454931
A:Accession: A28915
A:Molecule type: mRNA
A:Residues: 1-499,501-965, 'P', 967-1153 <ARN>
A:Cross-references: GB:M18044; GB:J03270; GB:M19664; GB:X07421; NID:G186935; PIDN:AAA5
A:Note: the authors translated the codon TAC for residue 1129 as Thr
A:Note: part of this sequence, including the amino end of the mature protein, was conf:
R:Shelley, C.S.; Arnaout, M.A.
Proc. Natl. Acad. Sci. U.S.A. 88, 10525-10529, 1991
A:Title: The promoter of the CD11b gene directs myeloid-specific and developmentally re
A:Reference number: A41600; MUID:92073318; PMID:1683702
A:Accession: A41600
A:Molecule type: DNA
A:Residues: 1-9 <SHE>
A:Cross-references: GB:M76724; NID:G180018; PIDN:AAA58410.1; PID:G553215
R:Arnaout, M.A.; Remold-O'Donnell, E.; Pierce, M.W.; Harris, P.; Tenen, D.G.
Proc. Natl. Acad. Sci. U.S.A. 85, 2776-2780, 1988
A:Title: Molecular cloning of the alpha-subunit of human and guinea pig leukocyte adhe
A:Reference number: A94193; MUID:88190151; PMID:2833753
A:Accession: A30892
A:Molecule type: mRNA
A:Residues: 917-1042 <AR2>
A:Cross-references: GB:M18044
R:Hickstein, D.D.; Hickey, M.J.; Ozols, J.; Baker, D.M.; Back, A.L.; Roth, G.J.
Proc. Natl. Acad. Sci. U.S.A. 86, 257-261, 1989
A:Title: cDNA sequence for the alphaM subunit of the human neutrophil adherence recept<
A:Reference number: A32218; MUID:89098893; PMID:2563162
A:Accession: A32218
A:Molecule type: mRNA
A:Residues: 9-1153 <HTC>
A:Cross-references: GB:J04145; NID:G189068; PIDN:AAA59503.1; PID:G386975

integrin - fruit f
integrin alpha-2b
integrin alpha-7 c
alpha-7 integrin -
glycoprotein IIB -
leukocyte adhesion
integrin alpha cha
F54F2.1 protein -
integrin alpha v c
hypothetical prote
integrin alpha cha
integrin alpha cha
glycoprotein IIB -
integrin alpha cha
glycoprotein IIB -
integrin alpha-1 -

A>Note: part of this sequence was confirmed by protein sequencing
 J. Fleming, J.C.; Pahl, H.L.; Gonzalez, D.A.; Smith, T.F.; Tenen, D.G.
 J. Immunol. 150, 480-490, 1993
 A>Title: Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-in
 n during evolution.
 A:Reference number: A46526; MUID:93123748; PMID:8419480
 A:Accession: A46526
 A>Status: not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-499,501-1153 <FILE>
 A:Cross-references: GI:S52227; NID:G263047; PIDN:AAB24821.1; PID:G263049
 A>Note: the last three bases of intron 13, CAG, are included in some but not all mature
 A>Note: sequence extracted from NCBI backbone (NCBIP:121963)
 R:Pierce, M.W.; Remold-O'Donnell, E.; Todd III, R.F.; Arnaout, M.A.
 Biochim. Biophys. Acta 874, 368-371, 1986
 A>Title: N-terminal sequence of human leukocyte glycoprotein Mol: conservation across sp
 A:Reference number: A90664; MUID:87076671; PMID:3539202
 A:Accession: A26091
 A:Molecule type: protein
 A:Residues: 17-31 <PIE>
 A:Experimental source: granulocytes
 R:Pahl, H.L.; Rosmarin, A.G.; Tenen, D.G.
 Blood 79, 865-870, 1992
 A>Title: Characterization of the myeloid-specific CD11b promoter.
 A:Reference number: 152567; MUID:92144986; PMID:1346576
 A:Accession: 152567
 A>Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-9 <RES>
 A:Cross-references: GB:M84477; NID:G180184; PIDN:AAA51960.1; PID:9553219
 A:Comment: A common beta chain (CD18) forms a heterodimer with this chain to form Mac-1
 C:Genetics:
 A:Gene: ITGAM; CR3A
 A:Cross-references: GDB:120599; OMIM:120980
 A:Map position: 16p11.2-16p11.2
 A>Note: promoter contains a GATA motif and two Sp1 consensus binding sites
 C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
 C:Keywords: alternative splicing; calcium; cell adhesion; glycoprotein; heterodimer; mag
 R:1-16/Domain: signal sequence #status predicted <SIG>
 R:17-1153/Product: cell surface glycoprotein CD11b #status experimental <MAT>
 R:17-1108/Domain: extracellular #status predicted <EXT>
 R:148-318/Domain: von Willebrand factor type A repeat homology <WA2>
 R:465-473/Region: calcium/magnesium binding #status predicted
 R:530-538/Region: calcium/magnesium binding #status predicted
 R:593-601/Region: calcium/magnesium binding #status predicted
 R:1109-1134/Domain: transmembrane #status predicted <TM>
 R:1135-1153/Domain: intracellular #status predicted <INT>
 R:85,240,391,469,693,697,735,802,881,901,912,941,947,979,994,1022,1045,1051,1076/Binding
 Query Match 99.3%; Score 5839; DB 1; Length 1153;
 Best Local Similarity 98.9%; Pred. No. 0;
 Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;
 Qy 1 FNLDTENAMTFQENARGFGQSVVQLQSGRVVVGAPQEIIVAAQNRGLYQCDYSTGSCPEI 60
 Db 17 FNLDTENAMTFQENARGFGQSVVQLQSGRVVVGAPQEIIVAAQNRGLYQCDYSTGSCPEI 76
 Qy 61 RLQVPEAVNMSLGLSLAATTSPPQLLAGPTVHTQCSNTYVKGCLFSGNLBQPOK 120
 Db 77 RLQVPEAVNMSLGLSLAATTSPPQLLAGPTVHTQCSNTYVKGCLFSGNLBQPOK 136
 Qy 121 FPARLRGCPQEDSDIAFLIDGSGSIIPHPFRMKELVSTIMEQLKSKTFLSIMQYSEEF 180
 Db 137 FPARLRGCPQEDSDIAFLIDGSGSIIPHPFRMKELVSTIMEQLKSKTFLSIMQYSEEF 196
 Qy 181 RIHFTKEFQNNPNSLKPITQLLGRHTATGLKRVRELFNTNGARKNAFKILLL 240
 Db 197 RIHFTKEFQNNPNSLKPITQLLGRHTATGLKRVRELFNTNGARKNAFKILVVI 256
 Qy 241 TDGEKFDPLGYEDVPIPELREGVIRYVLGFGDAFSEKSRQELNTVASKPPRDHVFQAN 300
 Db 257 TDGEKFDPLGYEDVPIPELREGVIRYVLGFGDAFSEKSRQELNTVASKPPRDHVFQVN 316

Qy 301 NFPAKTVQNLREKIFAIBGTQTGSSSPFHEMQBQFSAITNSGPIITVGSYDWAG 360
 Db 317 NFPAKTVQNLREKIFAIBGTQTGSSSPFHEMQBQFSAITNSGPIITVGSYDWAG 376
 Qy 361 GVFLYTSKESKTFINMTRVDSMDNAYLGYAAAILRNKVSQSLVLCAPRYOHIGLVAMFR 420
 Db 377 GVFLYTSKESKTFINMTRVDSMDNAYLGYAAAILRNKVSQSLVLCAPRYOHIGLVAMFR 436
 Qy 421 QNTGMWESNANVKGTOIGAYFGASICSVDVDSNGSTDLVLCAPHYIYQTRGGQSVVCP 480
 Db 437 QNTGMWESNANVKGTOIGAYFGASICSVDVDSNGSTDLVLCAPHYIYQTRGGQSVVCP 496
 Qy 481 PRQARWOCDAVLVGEQOPKRGFGAALTVLGDNVNGDKLTDVAICAPGEEDNRGAVLYLF 540
 Db 497 PRQARWOCDAVLVGEQOPKRGFGAALTVLGDNVNGDKLTDVAICAPGEEDNRGAVLYLF 556
 Qy 541 HGTSGSGIISPHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVLTWGAQGHVLLRSQ 600
 Db 557 HGTSGSGIISPHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVLTWGAQGHVLLRSQ 616
 Qy 601 PVLRVKAIMFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
 Db 617 PVLRVKAIMFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676
 Qy 661 YDLALDSGRPHSRVFNENKSTRQTVGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720
 Db 677 YDLALDSGRPHSRVFNENKSTRQTVGLTQTCETLKLQLPNCIEDPVSPIVLRNF 736
 Qy 721 SLVGTPLSAFGLNLRPVLAEDAQRLFTALFPFKKNCNDNI CODDLSITPSFMSLDCLVVG 780
 Db 737 SLVGTPLSAFGLNLRPVLAEDAQRLFTALFPFKKNCNDNI CODDLSITPSFMSLDCLVVG 796
 Qy 781 GRPEFNVTVVRNDGEDSYRTQVTFPFLDLSYRKVSTLQNRORSRRLACSSASSTEV 840
 Db 797 GRPEFNVTVVRNDGEDSYRTQVTFPFLDLSYRKVSTLQNRORSRRLACSSASSTEV 856
 Qy 841 SGALKSTCSINHPIPPENSEVTFTITFDVDSKASLGKLLKLVKANTSENNMPTNKTEP 900
 Db 857 SGALKSTCSINHPIPPENSEVTFTITFDVDSKASLGKLLKLVKANTSENNMPTNKTEP 916
 Qy 901 QLELPVKYAVVTVSHGVSTKYLNTASNTSRVQHQYQVSNIGORSLSPLISLVFLVPV 960
 Db 917 QLELPVKYAVVTVSHGVSTKYLNTASNTSRVQHQYQVSNIGORSLSPLISLVFLVPV 976
 Qy 961 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWVNCISIAVCQRIQCDIP 1020
 Db 977 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWVNCISIAVCQRIQCDIP 1036
 Qy 1021 FPGIQSEFNATLKGKLSPDVYIKTSHNLLIYSTAAILFNDSVFTLLPQCGAFVRSQTE 1080
 Db 1037 FPGIQSEFNATLKGKLSPDVYIKTSHNLLIYSTAAILFNDSVFTLLPQCGAFVRSQTE 1096
 Qy 1081 KVEPFVFPNPLPLIVGSSVGGILLALITAAALYKLGFFKQYKQKMMSSRGSGPPGAEPPQ 1137
 Db 1097 KVEPFVFPNPLPLIVGSSVGGILLALITAAALYKLGFFKQYKQKMMSSRGSGPPGAEPPQ 1153
 RESULT 2
 S00551
 leukocyte surface glycoprotein Mac-1 alpha chain precursor - mouse
 N:Alternate names: complement-3 receptor alpha chain
 C:Species: Mus musculus (house mouse)
 C>Date: 30-Sep-1989 #sequence_revision 30-Sep-1991 #text_change 22-Oct-1999
 C:Accession: S00551; I59078
 R:Pytela, R.
 EXBO J. 7, 1371-1379, 1988
 A>Title: Amino acid sequence of the murine Mac-1 alpha chain reveals homology with the
 A:Reference number: S00551; MUID:88312584; PMID:3044779
 A:Accession: S00551
 A:Molecule type: DNA
 A:Residues: 1-1153 <PVT>
 A:Cross-references: EMBL:X07640; NID:G52982; PIDN:CAA30479.1; PID:G52983
 A>Note: the authors translated the codon CAC for residue 569 as Gln

R:Sastre, L.; Roman, J.M.; Teplov, D.B.; Dreyer, W.J.; Gee, C.E.; Larson, R.S.; Roberts,
Proc. Natl. Acad. Sci. U.S.A. 83, 5644-5648, 1986
A:Title: A partial genomic DNA clone for the alpha subunit of the mouse complement recep
A:Reference number: 159078; MUID:86287312; PMID:2942940
A:Accession: 159078
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 11-44 <RES>
A:Cross-references: GB:M14293; NID:gl98993; PIDN:AAA39484.1; PID:g554193
C:Genetics:
C:Gene: Mac-1
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
C:Keywords: cell adhesion; glycoprotein; transmembrane protein
F:1-16/Domain: signal sequence #status predicted <SIG>
F:17-1153/Product: leukocyte surface glycoprotein Mac-1 alpha chain #status experimental
F:148-318/Domain: von Willebrand factor type A repeat homology <VWA>
F:1106-1129/Domain: transmembrane #status predicted <TM>

Query Match 75.6%; Score 4447; DB 2; Length 1153;
Best Local Similarity 73.5%; Pred. No. 3.8e-303;
Matches 836; Conservative 147; Mismatches 153; Indels 2; Gaps 2;

QY 1 FNLDTENAMTFQENARGFQSVVQLGSRVWVGAPOEIVAAQNRGLYOCXYSTGSCRP 60
DB 17 FNLDTEHMTFQENAKGRQNVVQLGTSVVAAPQEAQVAVQTGALYOCXYSTGRCHP 76
QY 61 RLQVPVEAVNMSLGLSLAANTSPOLLACPTVHQCSTENYVKGCLFLGSLNRQQPOK 120
DB 77 PLQVPEAVNMSLGLSLAVSTVPQQLACGPTVHONCKENTVNGLCYLFGLNLRPPQ 136
QY 121 FPEALRGCPQEDSDIAFLDGGSGIIPHDFRMLKLVSTIMEQLKSKTLFSLMQYSEEP 180
DB 137 FPEALRECPQEDSDIVFLDGGSGINIIDFQMKKEFVSTVMEQFKSKTLFSLMQYSEEP 196
QY 181 RIHFTFKBNPNRSLKIPITQLGRTHHTATGLKVVRELNFNITNGARKNAKILFLL 240
DB 197 RIHFTENDFKRNPSPSHSVPIQLNGRTKTASGIRKVVRELFKHNGARENAKILVVI 256
QY 241 TQGEKPGDPLGVDPVLPEDRGVIRYVLGFDAPSEKSRQELMTVASKPRDHVFOAN 300
DB 257 TQGEKPGDPLGVDPVLPEDRADRVIRYVLGFDAPSEKSRQELMTVASKPRDHVFOAN 316
QY 301 NFEALKTQVQNRKELFALTEGTQSGSSFEHMSQEGFSAATNGPLLTSTVGSVDWAG 360
DB 317 NFEALNTIQVQNRKELFALTEGTQSGSSFEHMSQEGFSAITNGPLLTSTVGSVDWAG 376
QY 361 GVPLTSEKSTFINNTRVDSMDAYLGYAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420
DB 377 GAFLYTSKDKVTFINTTRVDSMDAYLGYAAIILNRVQSLVGLGAPRYQHIGLVAMFR 436
QY 421 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYECTGCGVSVCPL 480
DB 437 ENFGTWEPHTSKGSGIYFGASLCSVDVDSNGSTDLVLIGAPHYEKTGCGVSVCPL 496
QY 481 PRGQARWQCDVAVLXGEQGPGRFCAALTVLGDVNGDKLTDVAICAPGEEDNRGAVYLF 540
DB 497 PRG-RARWQCEALLHGDQHPGRFCAALTVLGDVNGDKLTDVAICAPGEQENQAVYIF 555
QY 541 HGTSSGSIPTSHSQRISAGSLSPRLOYFQOSLSGGDLTMDGLVDLTGVAQGHVLLRSQ 600
DB 556 YGASIASASASHRIIGAHFSPGLQYFQOSLSGGDLTMDGLVDLTGVAQGHVLLRSQ 615
QY 601 PVLRYKALMEFNPVARNVFCNDQVVKGKAGEVRVCLHVQKSTRDLREGQIQSVVT 660
DB 616 PVLRLAETWTFSPKVARSVFACQGVLENKQDAGEVRVCLRVKTKDLREGDLSVTY 675
QY 661 YDLALDSGRPHSAVFNETNSTRGQTVLGLTQTCETLKLQPLNCIEDPVSPIVLRNLF 720
DB 676 YDLALDPVRSIRAFDEFKNNTRRTVFGLMQKCEKTLKLLPDCVDVDSVPIILRLNY 735
QY 721 SLVGTPLSAFGNLRVLAEDAQRFTALFPFRKNGCNDNI CODDSLSITFSFMSLCLVVG 780
DB 736 TLVGEPLRSFGNLRVPLANDAQRFMTAFPPFRKNGCNDNICODDSLSITWSAGLDTLVVG 795

RESULT 3
RWNUIC

cell surface glycoprotein CD11c precursor - human
N:Alternate names: leukocyte adhesion receptor p150,95 alpha chain
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1992; A35543; S00864
A:Accession: A35543; S00864
J:Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.
J. Biol. Chem. 265, 12750-12751, 1990
A:Reference number: A36584
A:Contexts: erratum
A:Accession: A36584
A:Molecule type: DNA
A:Residues: 1-1163 <COR>
A:Note: this revision to the sequence from reference A35543 includes the carboxyl end
R:Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.
J. Biol. Chem. 265, 2782-2788, 1990
A:Title: Genomic structure of an integrin alpha subunit, the leukocyte p150,95 molecule.
A:Reference number: A35543; MUID:90153906; PMID:2303426
A:Accession: A35543
A:Molecule type: DNA
A:Residues: 1-834 <CO2>
A:Note: this sequence has been revised in reference A36584
R:Corbi, A.L.; Miller, L.J.; O'Connor, K.; Larson, R.S.; Springer, T.A.
EMBO J. 6, 4023-4028, 1987
A:Title: cDNA cloning and complete primary structure of the alpha subunit of a leukocyte
A:Reference number: S00864; MUID:88166645; PMID:3327687
A:Accession: S00864
A:Molecule type: mRNA
A:Residues: 1-755, 'L', 757-1163 <CO3>
A:Cross-references: GB:M61695; EMBL:Y00093; NID:g487829; PIDN:AAA59180.1; PID:g487830
A:Note: part of this sequence was confirmed by protein sequencing
C:Comment: A common beta chain (CD18) forms a heterodimer with CD11b to form Mac-1 on my
C:Genetics:
A:Gene: GDB:ITGAX; CD11C
A:Cross-references: GDB:119758; OMIM:151510
A:Map position: 16p11.2-16p11.2
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
C:Keywords: calcium; cell adhesion; glycoprotein; heterodimer; magnesium; tandem repeat;
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-1163/Product: cell surface glycoprotein CD11c #status predicted <EXT>
F:20-1107/Domain: extracellular #status predicted <EXT>
F:149-319/Domain: von Willebrand factor type A repeat homology <VWA>
F:1108-1133/Domain: transmembrane #status predicted <TM>
F:1134-1163/Domain: intracellular #status predicted <INT>
F:61,89,392,697,735,899,939,1050/Binding site: carbohydrate (Asn) (covalent) #status pre

Db 1035 PPSVQSEEDFLTKGNLSGWRQILQKKVSVVAEITFDTSVYSQLPQGFMEAOQT 1094
 QY 1080 TKVEPPEVENPLIIVGSSVGGILLIALLIATAALYKLGFFKROVKDWMS 1128
 Db 1095 TVLEKYKVENPTPLIVGSSIGGLIALLIATAVLYKVGFFKROKEMWES 1143

RESULT 4
 S03308
 cell surface glycoprotein CD11a precursor - human
 C;Alternate names: leukocyte adhesion glycoprotein LFA-1 alpha chain; leukocyte
 C;Species: Homo sapiens (man)
 C;Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 20-Aug-1999
 C;Accession: S03308; A47458; A47565; A48759; S36044
 R;Larson, R.S.; Corbi, A.L.; Berhan, L.; Springer, T.
 J. Cell Biol. 108, 703-712, 1989
 A;Title: Primary structure of the leukocyte function-associated molecule-1 alpha
 A;Reference number: S03308; MUID:89139587; PMID:2537322
 A;Accession: S03308
 A;Molecule type: mRNA
 A;Residues: 1-1170 <LAR>
 A;Cross-references: EXBL:X00796; NID:g31421; PIDN:CRA68747.1; PID:g31422
 A;Note: part of this sequence was confirmed by protein sequencing
 R;Cornwell, R.D.; Gollahan, K.A.; Hickstein, D.D.
 Proc. Natl. Acad. Sci. U.S.A. 90, 4221-4225, 1993
 A;Title: Description of the leukocyte function-associated antigen 1 (LFA-1 or CD
 A;Reference number: A47458; MUID:93348261; PMID:8097887
 A;Accession: A47458
 A;Molecule type: DNA
 A;Residues: 1-20 <COR>
 A;Note: sequence extracted from NCBI backbone (NCBIN:130862, NCBI:P:130863)
 R;Shelley, C.S.; Farokhzad, O.C.; Arnaout, M.A.
 Proc. Natl. Acad. Sci. U.S.A. 90, 5364-5368, 1993
 A;Title: Identification of cell-specific and developmentally regulated nuclear f
 A;Reference number: A47565; MUID:93281759; PMID:8099450
 A;Accession: A47565
 A;Molecule type: DNA
 A;Residues: 1-20 <SHE>
 A;Cross-references: GB:M95609
 R;Nueda, A.; Lopez-Cabrera, M.; Vara, A.; Corbi, A.L.
 J. Biol. Chem. 268, 19305-19311, 1993
 A;Title: Characterization of the CD11a (alphaL, LFA-1alpha) integrin gene promot
 A;Reference number: A48759; MUID:93374910; PMID:8103515
 A;Accession: A48759
 A;Molecule type: DNA
 A;Residues: 1-20 <NUE>
 A;Cross-references: EMBL:Z22804; NID:g311405; PIDN:CAA80461.1; PID:g311406
 C;Genetics:
 A;Gene: GDB:ITGAL; CD11A
 A;Cross-references: GDB:119757; OMIM:153370
 A;Map position: 16p11.2-16p11.2
 C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A rep
 C;Keywords: cell adhesion; cytoskeleton; glycoprotein; heterodimer; surface anti
 P.1-25/Domain: signal sequence #status predicted <SIG>
 P.26-1170/Product: leukocyte adhesion glycoprotein LFA-1 alpha chain #status pres
 P.154-317/Domain: von Willebrand factor type A repeat homology <VWA2>

Query Match 26.1%; Score 1532.5; DB 2; Length 1170;
 Best Local Similarity 34.2%; Pred. No. 9,7e-99;
 Matches 401; Conservative 209; Mismatches 463; Indels 101; Gaps 37;

QY 1 FNLDTENMTQ--ENARGFGGSVVQLGSSRVVWVGAPOEIVANORGSLYQCDYSTGSC 58
 26 VNLDRGARSFSPPRAGHFGYRLQV--GNGVVGAPGE---GNSGTSLYQCQSGTGHCL 81
 59 PIRLOVPVEAVNMSLGLSIAATTSPPOLLACPTVHOICSENTVYKGLCFPLGSLNR--- 115
 82 PVTLR--GSNTYSKYLGMTLATDPTDGSILACDPGLSRICDQNTYLSGLCYLFRQNLQGM 140
 116 --QOPKQPFPEARLGCQPBDSIAFLIDGGSGIIPHDFRMKELIVSTIMBOLKKSKTLFSLM 174
 141 LQGRGFGFQECNG---NVDLVFLPDGSMISLQPDFKILDFMKDWKMLKSLNTSYQFAAV 196

175 QYSEPRIFHTKEFQNNPNSRLIKPITQILGRTHATGLKRVKVELFNITNGARKAP 234
197 QFSTSTYKTEFSDYVYKNDPALLKHVXHMILLTNTFGAINVATEVEREBLGRPDAT 256
235 KILFLITDEKFGDPLGYEDVPELDRGVIRYVGLGDAFSEKSRQSLNNTVASKPRPD 294
257 KVLIIITDGE--ATDSGNDAKD-----IIRYIIGIKGHFQKESQETLHKFASKPASE 309
295 HYFOANNFALKTVQNLREKFAHGTGTGSSSPHEHMSQSGFSAALTSNGPLLVTVG 354
310 FVKILDTPEKLDLFTLEOKKIVYIEGTSKDLTSFNLMELSSSGISADLSRGAHVAVG 369
355 SYDWAGGVF-LVTSKESKSPINNRVDSMDNDAIYGYAAA-IILNRVQSLVGLAPRYOH 412
370 AKDWAGGFLDLKADLQDDTFIGNEPITPEVRAGLYGVYTNWLPKQKTSLLASGAPRYOH 429
413 IGLVAMFR--QNTGMWESNANVKYQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQT 470
430 MGRVILLFQBPQGGHWSQVOTITGIGSYFGGELCGVDVDQDGETELLIGAPLYFGEQ 489
471 RGGQVSVCLPRGQARWCCDAV--LYGEGOPKGRFGAALTVLGDNVNGDKLTDVAIGAP 528
490 RGRVFIY-----QRQLGFEEVSELQGDGFLPLGRFGEAITALTIDINGDLVDVAVGAP 544
529 GEEDNRGAVYLFHGTSGSGISFSHQRIAGSKLSPQLQYFGOSLGGQDLTMDGLVDLV 588
545 LEE--QGAVYIENGHRG-GLSPQPSQRIEQTQVLSGIQWFGRSIHGWKLEGDLADVAV 601
589 GAQCHVLLARSOPVLAVKALMEFNEPREVARNVFECNDQVV-KKGAGEVRYVCLHVQKSTR 647
602 GAESQIVLUSRPVDMVTLMGFSFPAIIPVHEVECSYSTSNKMGWNITICFQI-KSLY 660
648 DLREGQIQSVVYTDIALDSGPHSRAVENETKNSRRQTVLGLTQTCETLKLQLPNCI 707
661 PQF-QGRVANLYTYTLQDGHRTREGLPPGGRHRLRNIAVT-TSMSCDTDFEPFVCV 718
708 EBPVSIVLRNPSL---VGTPLS--AFGN-----LRPVLADAQRLFTALPPFENKGN 757
719 QDLISPIVNSLFWEEBEGTRDQRAQKQIPIILRPSLHSETWEI-----PPEKNCE 773
758 DNICODLSITFSMSLDCLVGPPREFNVTVVANDGDSYRTQVTFPFLDLGVKVS 817
774 DKKEANLAVSPARSRLRLTAPASLSVLSLNLBEDAIVWQDLHLFPPLGSLFRKVE 833
818 TLQNRQSRWKLACES--ASSTEVSGALKTSYCSINHPIFENGEVNTNITFDVDSKAS 875
834 ML---KPHSQIPVSCHEEPESRLLSRAL---SCNVSSPIFRAGHSVALQMMFNTLVNSS 887
876 LGNKLALKANVTENN---MPRENTKEFQLELPVKYAVVMVYTSHGVSSTKYLNTASEN 931
888 WGSVELHANVTENNEDSLLEDNSATTI---IPILYPINILIQOEDSTLVSTPKGP 944
932 TSVVMQHOYQV---SNLGQSLP-ISLVLVPLVRLNQTVIMDRPQVTFSENLSSTCHTK- 986
945 KTHQVKMTQVRIQPSIHDHNIPTLEAVVGVQPPSEGGPIHQWSVQMBPPV--PCHVED 1002
987 -ELPSSHSD--FLAELRKAPVNCSTAVCQRIQCDDIPFGIEGEEFNATLKNLSFDWIK 1043
1003 LELPDAAPFCLPGALFRCPW-----FRQELVQVIGLELVGEIE 1044
1044 TSHNHLIVSTAEILFNDVFTLLPQCGAFVRSQETKVEPPEVENPLPLIVGSSVGGILL 1103
1045 AS-SMFLSCSLISISNSSKHPLHYSNLSL-AQVVMKVVDVYVERQMLVLYVLSGIGILL 1102
1104 LIALITRALKYKGFYKQYKDMMSBG-GPPGAP 1136
1103 LLLLIFIVLYKVGFFYKRLKMEAGRGVNGIP 1136

RESULT 5
156126
lymphocyte fuction-associated molecule-1-alpha - mouse

C:Species: Mus musculus (house mouse)
C:Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 20-Aug-1999
C:Accession: I56126
R:Kaufmann, Y.; Tseng, E.; Springer, T.A.
J. Immunol. 147, 369-374, 1991
A:Title: Cloning of the murine lymphocyte function-associated molecule-1 alpha-subunit
A:Reference number: 156126; MUID:91268576; PMID:2051027
A:Accession: I56126
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-1163 <RES>
A:Cross-references: GB:M60778; NID:g198785; PIDN:AAA39426.1; PID:g198786
C:Genetics:
A:Gene: LFA-1
A:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
P:151-315/Domain: von Willebrand factor type A repeat homology <VWA1>

Query Match 25.8%; Score 1516.5; DB 2; Length 1163;
Best Local Similarity 33.9%; Pred. No. 1.3e-97;
Matches 396; Conservative 216; Mismatches 466; Indels 89; Gaps 36;
QY 1 FNLDTEANMTFOENA-RGFGQSVVQLQGSRRVVGAPQEIIVAANQSGSLYQCDYSTGSCPE 59
DB 24 YNLDTRPTQSFQAQGRHFGVQLQIEDG-VVVGARGE---GDNTGGLYHCTSSSEFCQP 79
QY 60 IRLQVPVEAVMSLGLSLAATTSPPQLLACQPTVHQTCSNTYVKGCLFLPGSLMLRQPO 119
DB 80 VSLH-GSNHTSKYLGMTLATDAKGLSLIACDPLGSLRTCDQNTYLSGLCYLFPQSLGPM 138
QY 120 KFPENALRGCPQEDSDIAPLIDSGSIIPHDFRMRKELVSTIMEOLKSKTKLPSLMOYSEE 179
DB 139 QNRFAVQCMKGVLDVFLFDGQSOLDRKQDFEKLLEFMKQVRRKLSNTSYQFAAVQFSTD 198
QY 180 FRIHFTKEP-QNNPNRSLIKPITQLLGRTHATGLKRVKVELFNITNGARKAKILF 238
DB 199 CRTSFTFLDYVQKNKPNPDLVLLGSPQMPFLTNTFRAINVVAHVFKESGARPDATKVLV 258
QY 239 LITDGEKF--GDPLCYEDVIELEDRGVIRYVGLGDAFSEKSRQSLNNTVASKPRPDV 296
DB 259 IITDGEASDKGNISAAHD-----ITRYIIGIKGHFVSQKQTLHIIFASEVEFEV 309
QY 297 FOANNFALKTVQNLREKFAHGTGTGSSSPHEHMSQSGFSAALTSNGPLLVTVG 356
DB 310 KILDTPEKLDLFTLEOKKIVYIEGTSKDLTSFNLMELSSSGISADLSRGAHVAVGAP 369
QY 357 DWAGGVF-LYTSKESKSPINNRVDSMDNDAIYGYAAA-IILNRVQSLVGLAPRYOH 414
DB 370 DWAGGFLDLREDLQAGATPVGQEPILTSVDRGGYLGTYVAMWTSRSPPLLAACAPRYOHV 429
QY 415 LVAMFR--QNTGMWESNANVKYQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQT 472
DB 430 QVLIFQAPAGRNWQTKIEGTQIGSYFGGELCSVDLQDGEAELLIGAPLFPGEQ 489
QY 473 QGVSVCLPRGQARWCCDAVLYGEGOPKGRFGAALTVLGDNVNGDKLTDVAIGAPGED 532
DB 490 GRVFTY---QRRQSLFEMVSELQDGPVPLGRFGAAITALTIDINGDLTDVAVGAPLEE- 545
QY 533 NRGAVYLFHGTSGSGISFSHQRIAGSKLSPQLQYFGOSLGGQDLTMDGLVDLVTVGAQ 592
DB 546 -QGAVYIENGHRG-GLSPQPSQRIQGAQVFPGRFNGHSIHGVLDGLDRLADVVGAG 603
QY 593 HVLLRSQVPLRVKALMEFNEPREVARNVFECNDQVWKGKAG-BVRVCLHVQKSTRDLR 651
DB 604 RVVLSRPVVDVVTLSFSPEEIPVHEVECSYARBEQKHGKACFRKPLTPQ--F 661
QY 652 EGQTSQVVTYDLALDSGPHSRAVENETKNSRRQTVLGLTQTCETLKLQLPNCIEDPV 711
DB 662 QGRLLANLSYTLQDGHMRSGFLPPDGSBELSGNTSITP-DKSCLDHFHFHFIQIQILI 720
QY 712 SPIVLRNFSLV---GTPLSAFGN-LRPVLADAQRLFTALPPFENKGNINI QODLSI 767
DB 721 SPINVSLSLLEEGTPRDQKGRAMQPILRPSIHTV-TKEIPFENKGGCKKCEANLTL 779


```

QY 956 FLV-----PVELNOTVIWDRPOVTF-SERLS 980
D 985 LVTHIPIQYTKENKPLLYLTGQTDQAGDISCTAEINPLKLPHTA-----SVSFKNER 1040
QY 981 STCHTERKLSHSDFLAELRKAPVNVCSIAVCORICDIPFGIGIEFNATLKNLSFDW 1040
D 1041 ---HTKE-----LDCRTTSCSNITCWLKDLKAEYFYNVTVRWRT 1080
QY 1041 YIKTSHHLLIVTAETLNDVSFTLLPGOGAFVRSQTETKVPFVBNPLPLIVGSSVG 1100
D 1081 PAASTQTQVLTAAALIDTNPOLFVIEENAVTIPLMKPTKAEVPT--GVIIISIIA 1138
QY 1101 GLLLALITAAALYKLGFFKQYKDM 1125
D 1139 GILLALMTAGLWKLGFKKQYKDM 1163

RESULT 10
A33998
N;Integrin alpha-2 chain precursor - human
C;Species: Homo sapiens (man)
C;Date: 30-Mar-1990 #sequence revision 18-Sep-1992 #text_change 15-Sep-2003
C;Accession: A33998; B56793; A53117
R;Takada, Y.; Henler, M.E.
J. Cell Biol. 109, 397-407, 1989
A;Title: The primary structure of the VLA-2/collagen receptor alpha(2)-subunit (platelet)
A;Reference number: A33998; MUID:89308879; PMID:2545729
A;Accession: A33998
A;Molecule type: mRNA
A;Residues: 1-1181 <TAK>
A;Cross-references: GB:17033; NID:g33906; PIDX:CAA34894.1; PID:g33907
A;Note: the authors translated the codon GAT for residue 802 as Gin, GTC for residue 803
R;Cattamel, B.; Parmentier, S.; Leung, L.L.; McGregor, J.L.
Biochem. J. 279, 419-425, 1991
A;Title: Separation of important new platelet glycoproteins (GPIa, GPIC, GPIC*, GPIIa and
A;Reference number: A56793; MUID:92061944; PMID:1953640
A;Accession: B56793
A;Molecule type: protein
A;Residues: 30-43 <CAT>
A;Experimental source: platelet
R;Zutter, M.M.; Santoro, S.A.; Painter, A.S.; Tsung, Y.L.; Gafford, A.
J. Biol. Chem. 269, 463-469, 1994
A;Title: The human alpha-2 integrin gene promoter. Identification of positive and negative
A;Reference number: A53117; MUID:94103255; PMID:8276836
A;Accession: A53117
A;Molecule type: DNA
A;Residues: 1-16, 'V', 18-21 <ZUT>
A;Cross-references: GB:124121; NID:G900342; PIDX:AAA16619.2; PID:94583535
A;Note: authors translated the codon GTA for residue 17 as Leu
C;Genetics:
A;Gene: GDB:ITGA2; CD49B
A;Cross-references: GDB:128031; OMIM:192974
A;Map position: 5q11.1-5q11.2
C;Keywords: cell adhesion; glycoprotein; heterodimer; transmembrane protein
F:1-29/Domain: signal sequence #status predicted <SIG>
F:30-1133/Domain: extracellular #status predicted <EXT>
F:172-347/Domain: von Willebrand factor type A repeat homology <VWA2>
F:1134-1154/Domain: transmembrane #status predicted <TM>
F:1155-1181/Domain: intracellular #status predicted <CYT>
F:105,112,134,432,460,475,699,1057,1074,1081/Binding site: carbohydrate (Asn) (covalent)

Query Match 17.9%; Score 1054; DB 2; Length 1181;
Best Local Similarity 26.8%; Pred. No. 3.6e-65;
Matches 328; Conservative 215; Mismatches 494; Indels 188; Gaps 44;
QY 1 FNLDENATFTQ-ENARGGQSVQL-----QGRVVVGAPQEIIVANORGLYQC--DYST 54
D 30 YNVGPEAKIFGSPSEQGYAVQVQFIPKPKGNLLVGSFWSGPPENRMGVDVTKCPVDLST 89
QY 55 GSCPEIRLQ-----VPVEAVNMSLGLSLAATTSPQQLACGPTVHTCSENTVYKGLC 107
D 90 ATCEKLNLTSTISIPNVTETKNTMSLGLITRNMTGGFLTCGPLWAQCGNQYITTVGC 149

```

```

QY 108 FLFGNLRQOPKPFPEALRGCPQEDSDIAFLIDGSGSIIPHFRMKELVSTIMBOLK-- 165
D 150 SDISPDF-QLSASFSPATQCPSSL-IDVVVVCDENSIYFMD--AVKNLEKFXVQGLDIG 205
QY 166 KSKTLFSLMOYSEERLIHFTFKSPONPNPRSLIKPITCLLG-RHTHTATGLAKVRELEN 224
D 206 PKTQVGLIQYANRVRVFNLTNTYKEMIVATSTQSYGGDLNTFPATQIYARKYAS 265
QY 225 IYNGARKNAFKILFLTLDGEKFGDPLGYEDVTEIPDLREGVIRY---VLGF--GDAPRSK 279
D 266 AASGGRSATKVVVVVTDGESH-DGSMKXAVDQCNDHNLFGIAGVAVLYLNALDTKN 324
QY 280 SROELNTVASKEPRDRHVFQANNFEALKTQVQNLREKIFALEGTCTGSSSSPHEMSOEGF 339
D 325 LIKEIKATASIPTEYFFNVSDRAALLEKAGTLGEQIFSIETGVGQ-GDNFQEMSVQGF 383
QY 340 SAAITSNGP--LLSTVSGYDNAGGVPLVTSKEKSTFINMT--RVSDMN-DAYLGYAAAI 394
D 384 SADYSSQNDILMLGAVGAFGSGTIVQKTSHGHLIFPKQAFQILQDRNHSYLYGSVAA 443
QY 395 ILNRVQSLVILGAPRYQHILGVAMPQNTGWESNANV-----KGTQIGAYFGASLCSV 448
D 444 ISTGESTHFPVAGAPRANTYTGQIVLYSVN-----ENGNITVIAHHRGDQIGSYFGVLCV 498
QY 449 EYDSNGSTDLVLIGAPHYYEOTR--GGQVSVCPPLRGQARWQCDALVLYGEGQCPWGRFG 506
D 499 DVDKDTITDVLVAGAPMTSMDLKKEEGRVYLFTIKKILGQHQ---PLEGPRGIENTRFG 555
QY 507 AALTVLGVNKGDLTDVAIGARDEEDREGAVLYLFGHTSGSGISPSHSRIAGS--KLSPR 564
D 556 SAIAALSINNDGFDVIVGSPLENQNSGAVIYVNGHQGT-IRTKYSQIKLSDGDAFRSH 614
QY 565 LQYFOQSISGGQDITMDCLVLTGVAQGHVLLLSQPVLRKAIMFNPREFVARVNFBCN 624
D 615 LQYFGRSLDYGDLNGSDITDVSIGAFQVQVQLMSQSIADVAIEASFPEKI--TLVNKN 672
QY 625 DOVVKKEAGEVRVCLHVQKSTRDLREGQIQSVVYDIALD----SGRPHSAVFNETK 680
D 673 AQII-----LKLCF-----SAKFRPTQNNQVAIVYNTILDAGFSRVRTSRGLFKENN 721
QY 681 NSTRRQTVLGLTQTC--ETLKLQLPNCIEDVPSPIVLRNLFNPLVGLPAGNLRPLVIA 738
D 722 ERLCKNVVNVQASCPHHIYIIEPS---DVVNSLDLREVISLENPGTS-----PALE 772
QY 739 EDAQRLFTALPFRKNCNDNICQDDLSITF-----SFMSLDCLVVGGRPFNVTVTRND 794
D 773 AYSETAKVFSIPFKHCKGEGDGLCISDLVDVRQIPAAQEQPFIVSNQNKRLTFSVTLXNK 832
QY 795 GEDSVRTQVTFPPPLDLSYRKVSTLQNSORSRSHLACESAST-EVSGALKSTSCSINH 853
D 833 RESAINTGIVVDFSENLPF-----ASFSLPVDGTETVTCQVAASQKSCVADGVY 880
QY 854 PIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPRTNKTETFOLELPVYAVTMV 913
D 881 PALKREQVTTTFINFDPLQ-NLQVQASISFOALSESQENKADNLVNLKIPLLYDAEI- 938
QY 914 VTHGVSTKYLNPTASENTSVMQHVQVSNLGR-----SLPISLVFLV----- 958
D 939 ---HLTRSTNINPFEISDGNVPSIVSHSFEDVGPKPIFSLKVTGSGVPSMATVTHIHP 995
QY 959 -----PVRLAQTVIWDPRQVTF-SENLSSTCTHKEK 988
D 996 YTEKNPLMYLTGVQTDKAGDISCNADINPLKIQGT---SSSVSPKSENF--HTKE- 1047
QY 989 LPSSHDFLAELRKAPVNVCSIAVCORIQCDIPFGIGIEFNATLKNLSFDWYIKTSNH 1048
D 1048 -----LNCRTASCNTVCKLKVHMKGEYFNVVTVTRINWTPASSTFQT 1091
QY 1049 LLIVSTAEI-LPNDVSPTLLPGOGAFVRSOTETKVPFVBNP-----LP--LVGSSVG 1100
D 1092 VQLTAAAEINTYNPISYVI-----EDNTVTIPLMINKPDEKAEVPGVIGLSIIA 1141

```


1101 GULLLALITAAALYKLGFFKQYKDM 1125
 1142 GULLLALVAILAKLGFFKQYKDM 1166

RESULT 11
 A35854
 integrin alpha-1 chain precursor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 23-Oct-1990 #sequence_revision 13-Sep-1991 #text_change 15-Sep-2003
 C:Accession: A35854; S11243
 R:Ignatius, M.J.; Large, T.H.; Houde, M.; Tawil, J.W.; Barton, A.; Esch, F.; Carbonetto, J. Cell Biol. 111, 709-720, 1990
 A:Title: Molecular cloning of the rat integrin alpha-1-subunit: a receptor for laminin A
 A:Reference number: A35854; MUID:90338125; PMID:2380249
 A:Accession: A35854
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1180 <IGN>
 A:Cross-references: GH:X52140; NID:G56493; PIDN:CAA36384.1; PID:G56494
 C:Keywords: cell adhesion; cytoskeleton; transmembrane protein
 F:170-345/Domain: von Willebrand factor type A repeat homology <VWA2>

Query Match 17.8%; Score 1049; DB 2; Length 1180;
 Best Local Similarity 27.3%; Pred. No. 8e-65;
 Matches 344; Conservative 200; Mismatches 476; Indels 240; Gaps 48;

Qy 1 ENLDTENAMTQENARG-FGQSVVOL---QGSVVVQAPQEIIVANORGLYQCDYSTGS 56
 Db 29 FNVQVNSMSPGVEDMFGYTVQYENEEGKWLIGSLVGPQKARTGDVYKCPVGRER 88
 Qy 57 CEP-IRLQVPEA-----VNMSLGLSLAATTSPPQLLACQPTVHQTCSENVTVKGL 106
 Db 89 AMPCVKLDLPVNTSIPNVTEIKENNTTGSTL-VTNPNNGFLACPLVAYRCHLHYTGI 147
 Qy 107 CFLPESNLROQPKFPALRGCPQEDSDIAFLDGSGLIIPHPRMKEIVSTIMEOLKK 166
 Db 148 CSDVSPFQVNSAP--VQBCSTQ-LDVIIVLDGNSIYP-----WESVIAFLNDLKR 199
 Qy 167 -----SKTLFSLMQYSEFRHFTFEKQNNPNRSLIKPTQLLG-RHTATGLRKVR 220
 Db 200 MDIGPKQTVGVQYGVNTHFEMLNKYSSTEEVLVANKIGROGGLQTVTALGIDTARK 259
 Qy 221 ELFNITNGARKNAFKILFLTDGEGFGDPLGYEDVPELDRGVIRY---VLGFGDA--F 275
 Db 260 EAFTFARGARGVKKVWVIVTDGESH-DNYELKQVIOQCEDENIORPSIALGHYRGNL 318
 Qy 276 RSEKSRQELNTVASKPRADHVPQANNFEALKTQVQNLREKIFATECTQTGSSSFEHMS 335
 Db 319 STERFVEIKSIASEPTEKHFVNSDELALVTIVKALGERIFALEATDAQSAFEMEXS 378
 Qy 336 QEGPSAATISNGLISTVGSYDMAGVFLYTSKEKSTFINMT--RVDSMDND---AYLGY 390
 Db 379 QTGFSAHYQDQWMLGAVGADNGTVVMQKQANQWVPHNTTPTQTEPAQWNEPLASYLYG 438
 Qy 391 AAAILNRVQSLVGLAPRYQHIGLVAMFRONTGWSNANVKGTQIGATPGASLCSDVD 450
 Db 439 TVNSATIPGDVLYTAGQRYNHTQWVIYKQMEDGNINILQTLGGEQIGSYFGSVLYATIDI 498
 Qy 451 DSNGSTDLAVLIGAPHY-----YEQTR-GGQVSVCPPLRGQARQCCDAVL 494
 Db 499 DKQSTYDILLVGPAMTNGTEBQKQVTVYAVNQTRFEYQMSLEPIRQTCCSSLKDNST 558
 Qy 495 YGEGQPMG-RFGAALTVLGVNGKLTDAVAGPGBEDNKGAVYLPFGTSGGSISSH 553
 Db 559 KENKNEPCGAFGTAAAKVNDGVDVIGAPLEDHAGAVYIYEH-SGKTIREAYA 617
 Qy 554 QRIAGSKLSPLQVQSGSLGQDITMDGLVDLTVGAGCHVLLRSQPLRVKAIMFN 613
 Db 618 QRIQSGGQKTLKFPQSGIHGEMDNLGDLTDVTIIGLGGAALEFARDVAVKTMFEP 677
 Qy 614 REVARNVFECNDQVVKGEAG--EVRVCLHVQ-KSTRDLRLBEGQIQSVTVYDLALDGRP 670

678 NKVNQKNCNCR---VEGKETVCINATWCFHVKLSKEDSIYEADLQ-----YRVTLDSLRQ 730
 671 HSRAYFNET-----KNSTRQTVGLTQTCETLKLQPLNCI-----EDPVSPIV 715
 731 ISRSFSGTQERKIQRNITRESE-----CIRHSFYMLDKHDPQDSVR 773
 716 LELNFSLVGTPLSAPGNLRLPVLAEQRLFTALPPEKNCNDNQCQDDLSITPFSMSLD 775
 774 VTLDFNLT-DPENG-----PVLDDALPNSVHEHIFPACDCKGKRCISDLTLNVTTEKS 827
 776 CLVVGGPPE-FNVTVTVRNDGEDSYRTQVTFPPFLDLSTYRKVSTLQONQORSQSWLACES 834
 828 LLIVKSHDKKNSLTVKNGDSAYNTRVWQHSNPLIFSGIEEIQD-----SCBS 879
 835 ASSTVSGALKSTCSINHPFPENSEVTFNITVDYDSKASLGNKLL-LKANVTSENMP 893
 880 N-----QNTICRVGYPPLRAGETVTFKIIFQFNTSHLSENALIHLSATSDSEEPLE 930
 894 RTNKTPEOLELPPVKYAV---YMWVTSHGVSF-----KYNLFTASENITSVMQYQV 942
 931 SLNDNEVNISIPVKEVGLQFVSSASEHHISVAANETIPEFINST--EDIGNEINVFTI 988
 943 SNLQORSLP---ISLVF-----LVPVRLNQTIVD-----RP----- 971
 989 RRGHPMPPELQLSISFPNLTADGYFVLPYG-----WSSDNVNCRRSLEDPPGINS 1042
 972 --QVTFG-----ENLSSTCHTKERLPKSHSDFLAELRKAPVNVVCSIAVCORIQCDI 1019
 1043 GKMTIISKEVLKRTGIQDCSSTC-----GVANTICSLPSDLSQVNV 1085
 1020 -----PFGIQEFP---NATLKNLSPDWYIKTSHNHLIVSTAEILFNDVSFTLLPQ 1070
 1086 SULLMKPTE-IRAHFSSSLNLTIRGELK-----SENSSLTSSN----- 1123
 1071 GAFVRSQTEKVEPEVENPLPL--IVGSSVGLLLALITAAALYKLGFFKQYKDMMS 1128
 1124 ---RKELAIQIKQGLPCRVPLWILLAPAGLLALLLALWKLKLGFFKPKKKKKEK 1180

RESULT 12
 A41131
 lymphocyte-Peyer's patch adhesion molecule alpha 4 precursor - mouse
 N:Alternate names: integrin alpha-4
 C:Species: Mus musculus (house mouse)
 C:Date: 05-Jun-1992 #sequence_revision 05-Jun-1992 #text_change 24-Sep-1999
 C:Accession: A41131; S16742
 R:Neuhaus, H.; Hu, M.C.T.; Hemler, M.E.; Takada, Y.; Holzmann, B.; Weissman, I.L.
 J. Cell Biol. 115, 1149-1158, 1991
 A:Title: Cloning and expression of cDNAs for the alpha subunit of the murine lymphocyte
 A:Reference number: A41131; MUID:92064645; PMID:1840602
 A:Accession: A41131
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1039 <NEU>
 A:Cross-references: EMBL:X53176; NID:G51484; PIDN:CAA37316.1; PID:G51485
 C:Superfamily: integrin alpha-4 chain
 C:Keywords: cytoskeleton; transmembrane protein

Query Match 11.3%; Score 666; DB 2; Length 1039;
 Best Local Similarity 22.8%; Pred. No. 4.6e-38;
 Matches 283; Conservative 183; Mismatches 405; Indels 368; Gaps 48;

Qy 1 ENLDTENAMTQ-ENARGFGQSVV-QLOGSR--VVVGAPOEIVAAV---QRGSLYQCDY 52
 Db 41 YNLDPENALLYQGPSGTLFGYSVVLHSGSKRWLVGAPTASWLSNVSNNVPCALYRCGI 100
 Qy 53 STG---SCSPIRLQVP-----VEAVNMSLGLSLAATTSPP-PQLLACG---PTVHQ 95
 Db 101 RKNPNQTCQLOGSGSPGPGTKCLERDQNLVTLRSQPGENGSIIVTCGHRWNKIFY 160
 Qy 96 TCSNTYVYKGLCLFGLSNLRQQPKPFEALRGCPQEDSDIAFLDGSGLIIPHDFFRWE 155
 Db 161 MKSDNKLPTGICVMPSDRLTELK-----RMAP 189

Db 963 TVVFEALHLEPRGYVVGWIIAISLLVGLIFLLAVLAKMGFFRRRYKEII 1015

RESULT 15

T31437

integrin alpha chain SU2 - sea urchin (lytechinus variegatus)

C:Species: lytechinus variegatus (variegated urchin)

C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jan-2000

C:Accession: T31437

R:Hertzler, P.L.; McClay, D.R.

submitted to the EMBL data Library, May 1998

A:Description: Alpha SU2, a sea urchin integrin which binds laminin.

A:Reference number: Z21035

A:Accession: T31437

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1041 <R>

A:Cross-references: EMBL:AF067658; NID:g3220240; PID:g3220241; PIDN:AAC23572.1

A:Experimental source: developmental stage embryo

C:Function:

C:Description: binds laminin

C:Superfamily: integrin alpha-2b chain

Query Match 10.5%; Score 614.5; DB 2; Length 1041;

Best Local Similarity 24.8%; Pred. No. 1.9e-34;

Matches 236; Conservative 155; Mismatches 314; Indels 245; Gaps 41;

QY 336 QEGFSAATISNGP--LLSTVGSYDWAGVGLYTSKEKSFNNTRVDS-----DMNDAYLG 389

Db 181 QAGFSGILFSDNSALVMGAPGSYYLQGIYVQSLNRSV-VQATQESNTGYSPDNSYRG 239

QY 390 YAAAL--ILNRVQSLVLCAPRYQHI-GLVAFPRONTGMESNANVKGTQIGAYFGASLC 446

Db 240 YSLALGDFNGDGVDQYVVGTPRAESLMGLVAIFDQNLQFN---QVMGTQIVAYFGYSVT 296

QY 447 SYVDVNSGTDLVLIGAPHYVETRGGQVSVCPFRGQARQCDAVLVYGEQ----- 498

Db 297 VVDI-NDYDILLVGAPEYMDGPAIQ-----RWEAGAVYYLQNPVVGPGA 343

QY 499 -----GQWGRFGAALTVLGVNGDKLTDVAIGAPGEDNKGAVYLFHGTSGS 546

Db 344 SNRLSLSLTLGGQIRSRFGLSIASIGDSNQDFNDVAIGAPYEGDDAGAVYVHG-SAN 402

QY 547 GISPSSHQRIAGKLS-PRLOYFGOSLGGDLTDGLVDLTVGAQ--GHVILLRSQVPL 603

Db 403 GLKSTPAQVLTSTLGHSGITTFGSLQGGQMDKKNYFDLLVGAESANTAVLINTREPV 462

QY 604 RYKAIMFNPREVARNFECNDQVYVKGAGEVRVCLHVQKSTRDLREGQIQSV----- 658

Db 463 SLDATLNTPE-----IGINLENKTYE-LADGTWTSFIAMT 497

QY 659 -----VTYDLALDSG-RPHSRVFNTRKSTRQTVGLG-TOTCETLKLQ 702

Db 498 CFTYTGNYLPDHIDISYTVYVDSGHIANRRAMFVNDMSEITKTRRLAVSTQPCDPLRAY 557

QY 703 LPNCIEDPVSPVLRNLNLSVGTPLSAFCN-----LRPVLAEADAQLFTALFPPEK 753

Db 558 VGNSTIEDKLTPIKVTLYQYDL-----NNDESRLQPHLEIPIIDMATMSTQTKQVSIQN 609

QY 754 NCGNDNIQDDLSITFSPMSGLDCLVGGPRENVTVTNRNDEDSYRTQVTFPPPLDLSY 813

Db 610 KCVN-NICIPDLQVTVT-PNLFNIVIGQTELTDVSLNRRGEDAFQSSLSVYPLGLQF 667

QY 814 RYKSTLQHQRSQRSLACESASSTEVSGALKSTCSINHPIFFEN-----SEVTENIT 867

Db 668 VPL-----ERKANMDFSVTCSEDS-----LRITCTGNPMVGNKILFGLTSLTFQVS 717

QY 868 FVDVS-----KASLGNKLLKANVTSENNMPTNKTEFQLELPVKYAVYVMTVSHGVSTK 922

Db 718 GPKDSIEFFYKAESENS--EDPNTLNNELNMTVPVVDCTLKLSASYPEIMVMTQED 775

QY 923 YL-----NFTASENTSRVMCHOYQV-----SNLQESLPS-----LVFLVPV 960

Db 776 YVPPFPKANSEADIGNEVM-HLYEVRNTGSSNAGEVSLNIQWPKNEDGEYLFYLLGI 834

QY 961 RLNOTVIMDRPQ-----VTFSENLSSTCHTKERLPFSHSDFTLAELRKAPVV 1005

Db 835 MTEEGVTCQLTQGGKANPEGVKLEPSTKAKLSNSTTOVSGRKRREFEVAEALAQTDN--VI 892

QY 1006 NCSTAVCQRIQCDIPFFGCIQEEFNAT-----LKGNSLFDWYIKTSHHLLIVSTABILF 1059

Db 893 YCASDSCVLINCTI-----DEINASKSKVVRILGRP---W-----ERTF 928

QY 1060 NDSVFTLLPGGQAFVRSTQETKVB--PFEVNP-----LP----- 1092

Db 929 QKAVSELTPVQVQATIASASAAVXTIPYNIPLPRDFSDSTKASTLVITTEELVPPVPTIAW 988

QY 1093 -LIVGSSVGGELLALALTAALYKLGPFKR-----QYKDMMS--EGGPP 1132

Db 989 WIIVSVLGGIILLIILGLWKCGPFERKPKGEEKYAPVASADKDGPP 1038

Search completed: June 7, 2004, 17:17:38

Job time : 19.559 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 16:55:38 ; Search time 9.97194 Seconds
(without alignments)
5937.039 Million cell updates/sec

Title: US-09-902-481b-3

Perfect score: 5879
Sequence: 1 FNLDTENAMTFQENARGFQ.....FKRQYKDMSEGGPPGAEPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5823.5	99.1	1152	1	ITAM HUMAN
2	4447	75.6	1153	1	P05555 mus musculus
3	3446	58.6	1163	1	ITAD HUMAN
4	3388	57.6	1162	1	ITAD HUMAN
5	1536.5	26.1	1170	1	ITAD HUMAN
6	1116.5	25.8	1163	1	ITAD_MOUSE
7	1136.5	19.3	1167	1	ITAD_MOUSE
8	1128	19.2	1179	1	ITAE HUMAN
9	1085.5	18.5	1151	1	ITAI HUMAN
10	1079.5	18.4	1189	1	ITAH HUMAN
11	1071	18.2	1170	1	ITAB_BOVIN
12	1057	18.0	1178	1	ITAB_MOUSE
13	1054	17.9	1181	1	ITAB_MOUSE
14	1049	17.8	1180	1	ITAI RAT
15	1042.5	17.7	1167	1	ITAG HUMAN
16	666	11.3	1039	1	ITAA_MOUSE
17	642	10.9	1038	1	ITAA HUMAN
18	630	10.7	1035	1	ITA9 HUMAN
19	594.5	10.1	1032	1	ITAA_XENLA
20	571.5	9.7	1066	1	ITAC_CRISP
21	567.5	9.7	1053	1	ITAB_MOUSE
22	555.5	9.4	1053	1	ITAS_MOUSE
23	551.5	9.4	1034	1	ITAV CHICK
24	542.5	9.2	1050	1	ITAS_XENLA
25	539	9.2	1044	1	ITAV_MOUSE
26	534.5	9.1	1130	1	ITAE HUMAN
27	533.5	9.1	1072	1	ITAC_CHICK
28	532	9.0	1049	1	ITAS_HUMAN
29	531.5	9.0	1066	1	ITAB_HUMAN
30	530	9.0	1048	1	ITAV_HUMAN
31	514.5	8.8	1044	1	ITAB_CHICK
32	511.5	8.7	1091	1	ITAE_MOUSE
33	505	8.6	1179	1	ITAT_MOUSE

ALIGNMENTS

RESULT 1

ID	ITAM HUMAN	STANDARD;	PRT;	1152 AA.
AC	P11215;			
DT	01-JUL-1989 (Rel. 11, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha subunit) (CR-3 alpha chain) (CD11b) (leukocyte adhesion receptor MO1)			
DE	(Neutrophil adherence receptor).			
GN	ITGAM OR CR3A OR CD11B.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88315033; PubMed=2457584;			
RA	Corbi A.L., Kishimoto T.K., Miller L.J., Springer T.A.;			
RT	"The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CD11b) alpha subunit. Cloning, primary structure, and relation to the integrins, von Willebrand factor and factor B.";			
RL	J. Biol. Chem. 263:12403-12411(1988).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88190151; PubMed=2833753;			
RA	Arnsaut M.A., Remold-O'Donnell E., Pierce M.W., Harris P., Tenen D.G.;			
RT	"Molecular cloning of the alpha subunit of human and guinea pig leukocyte adhesion glycoprotein Mo1: chromosomal localization and homology to the alpha subunits of integrins.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 85:2776-2780(1988).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88257215; PubMed=2454931;			
RA	Arnsaut M.A., Gupta S.K., Pierce M.W., Tenen D.G.;			
RT	"Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor Mo1 (complement receptor type 3).";			
RL	J. Cell Biol. 106:2153-2158(1988).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=93123748; PubMed=8419480;			
RA	Fleming J.C., Pahl H.L., Gonzalez D.A., Smith T.F., Tenen D.G.;			
RT	"Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-integrin gene family demonstrate remarkable conservation of genomic organization and suggest early diversification during evolution.";			
RL	J. Immunol. 150:480-490(1993).			
RN	[5]			
RP	SEQUENCE OF 9-1153 FROM N.A.			
RX	MEDLINE=89098893; PubMed=2563162;			
RA	Hickstein D.D., Hickey M.J., Ozols J., Baker D.M., Back A.L., Roth G.J.;			
RT	"cDNA sequence for the alpha M subunit of the human neutrophil adherence receptor indicates homology to integrin alpha subunits.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 86:257-261(1989).			
RN	[6]			

34	497	8.5	1396	1	ITA2_DROME
35	491.5	8.4	1146	1	ITAB_DROME
36	490	8.3	1039	1	ITAB_HUMAN
37	489	8.3	1025	1	ITAB_MOUSE
38	488	8.3	1033	1	ITAB_MOUSE
39	486	8.3	126	1	ITAM_CAVPO
40	478	8.1	1181	1	ITAB_HUMAN
41	470	8.0	1106	1	ITAT_RAT
42	466	7.9	1226	1	ITAT_CAEEL
43	444.5	7.6	1139	1	ITAB_MOUSE
44	419	7.1	1115	1	ITA3_DROME
45	382	6.5	1000	1	ITA5_DROME

P12080	drosophila
Q24247	drosophila
P08514	homo sapien
P53708	homo sapien
Q9qum0	mus musculus
P11578	cavia porce
Q13583	homo sapien
Q63258	rattus norv
P34446	caenorhabdi
Q03600	caenorhabdi
O44386	drosophila
Q9wim8	drosophila

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J03925; AAA59544.1; -
CC EMBL; M18044; AAA59491.1; -
CC EMBL; J04145; AAA59903.1; -
CC EMBL; S52227; AAB24821.1; -
CC EMBL; S52152; AAB24821.1; JOINED.
CC EMBL; S52153; AAB24821.1; JOINED.
CC EMBL; S52154; AAB24821.1; JOINED.
CC EMBL; S52155; AAB24821.1; JOINED.
CC EMBL; S52157; AAB24821.1; JOINED.
CC EMBL; S52159; AAB24821.1; JOINED.
CC EMBL; S52161; AAB24821.1; JOINED.
CC EMBL; S52164; AAB24821.1; JOINED.
CC EMBL; S52165; AAB24821.1; JOINED.
CC EMBL; S52167; AAB24821.1; JOINED.
CC EMBL; S52169; AAB24821.1; JOINED.
CC EMBL; S52170; AAB24821.1; JOINED.
CC EMBL; S52173; AAB24821.1; JOINED.
CC EMBL; S52174; AAB24821.1; JOINED.
CC EMBL; S52180; AAB24821.1; JOINED.
CC EMBL; S52181; AAB24821.1; JOINED.
CC EMBL; S52184; AAB24821.1; JOINED.
CC EMBL; S52189; AAB24821.1; JOINED.
CC EMBL; S52191; AAB24821.1; JOINED.
CC EMBL; S52192; AAB24821.1; JOINED.
CC EMBL; S52203; AAB24821.1; JOINED.
CC EMBL; S52212; AAB24821.1; JOINED.
CC EMBL; S52213; AAB24821.1; JOINED.
CC EMBL; S52216; AAB24821.1; JOINED.
CC EMBL; S52219; AAB24821.1; JOINED.
CC EMBL; S52220; AAB24821.1; JOINED.
CC EMBL; S52221; AAB24821.1; JOINED.
CC EMBL; S52222; AAB24821.1; JOINED.
CC EMBL; S52226; AAB24821.1; JOINED.
CC EMBL; M76724; AAA58410.1; -
CC EMBL; M8477; AAA51960.1; -
CC PIR; A31108; RWHULB.
CC PDB; 1ABX; 17-JUN-98.
CC PDB; 1BHO; 18-NOV-98.
CC PDB; 1BHQ; 18-NOV-98.
CC PDB; 1IDN; 25-NOV-98.
CC PDB; 1IDO; 01-AUG-96.
CC PDB; 1JLM; 11-JAN-97.
CC PDB; 1MLU; 07-AUG-02.
CC Genew; HGNC:6149; ITGAM.
CC MIM; 120980; -
CC GO; GO:0008305; C:integrin complex; TAS.
CC GO; GO:0004895; F:cell adhesion receptor activity; TAS.
CC GO; GO:0007155; P:cell adhesion; TAS.
CC InterPro; IPR000413; Integrin_alpha.
CC InterPro; IPR02035; VWF_A.
CC Pfam; PF01839; FG-GAP; 3.
CC Pfam; PF00357; integrin_A; 1.
CC Pfam; PF00092; vwa; 1.
CC PRINTS; PR01185; INTEGRINA.
CC PRINTS; PR00453; VWFADOMAIN.
CC SMART; SM00191; Int_alpha; 4.
CC SMART; SM00327; VWA; 1.
CC PROSITE; PS00242; INTEGRIN_ALPHA; 1.
CC PROSITE; PS02344; VWFA; 1.
CC Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
CC Signal; 3D-structure; Repeat; Magnesium; Calcium.
CC SIGNAL 1 16
CC FT CHAIN 17 1152 INTEGRIN ALPHA-M.
CC FT DOMAIN 17 1104 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 1105 1128 POTENTIAL.
CC FT DOMAIN 1129 1152 CYTOPLASMIC (POTENTIAL).
CC -----

RP SEQUENCE OF 1-9 FROM N.A.
RX MEDLINE=92073318; PubMed=1683702;
RA Shelley C.S., Arnaout M.A.;
RT "The promoter of the CD11b gene directs myeloid-specific and
RT developmentally regulated expression.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:10525-10529(1991).
RN [7].
RP SEQUENCE OF 1-9 FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=92144986; PubMed=1346576;
RA Pahl H.L., Rosmarin A.G., Tenen D.G.;
RT "Characterization of the myeloid-specific CD11b promoter.";
RL Blood 79:865-870(1992).
RN [8].
RP SEQUENCE OF 17-31.
RX MEDLINE=87076671; PubMed=3539202;
RA Pierce M.W., Remold-O'Donnell E., Todd R.F. III, Arnaout M.A.;
RT "N-terminal sequence of human leukocyte glycoprotein Mol:
RT conservation across species and homology to platelet IIB/IIIA.";
RL Biochim. Biophys. Acta 874:368-371(1986).
RN [9].
RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 148-331.
RX MEDLINE=95171458; PubMed=7867070;
RA Lee J.O., Rieu P., Arnaout M.A., Liddington R.;
RT "Crystal structure of the A domain from the alpha subunit of integrin
RT CR3 (CD11b/CD18).";
RL Cell 80:631-638(1995).
RN [10].
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 148-334.
RX MEDLINE=96363671; PubMed=8747460;
RA Lee J.O., Bankston L.A., Arnaout M.A., Liddington R.C.;
RT "Two conformations of the integrin A-domain (I-domain): a pathway for
RT activation?";
RL Structure 3:1333-1340(1995).
RN [11].
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 148-337.
RX MEDLINE=98362595; PubMed=9687375;
RA Baldwin E.T., Sarver R.W., Bryant G.L. Jr., Curry K.A.,
RA Fairbanks M.B., Finzel B.C., Garlick R.L., Heinrichson R.L.,
RA Horton N.C., Kelley L.L., Mildner A.M., Moon J.B., Mott J.E.,
RA Mutchler V.T., Tomich C.S., Watenpaugh K.D., Wiley V.H.;
RT "Cation binding to the integrin CD11b I domain and activation model
RT assessment.";
RL Structure 6:923-935(1998).
RN [12].
RP 3D-STRUCTURE MODELING OF 17-616.
RX MEDLINE=98226734; PubMed=9560195;
RA Ovig C., Springer T.A.;
RT "Experimental support for a beta-propeller domain in integrin alpha-
RT subunits and a calcium binding site on its lower surface.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:4870-4875(1998).
CC -!- FUNCTION: INTEGRIN ALPHA-M/BETA-2 IS IMPLICATED IN VARIOUS
CC ADHESIVE INTERACTIONS OF MONOCYTES, MACROPHAGES AND GRANULOCYTES
CC AS WELL AS IN MEDIATING THE UPTAKE OF COMPLEMENT-COATED PARTICLES.
CC IT IS IDENTICAL WITH CR-3, THE RECEPTOR FOR THE IC3B FRAGMENT OF
CC THE THIRD COMPLEMENT COMPONENT. IT PROBABLY RECOGNIZES THE R-G-D
CC PEPTIDE IN C3B. INTEGRIN ALPHA-M/BETA-2 IS ALSO A RECEPTOR FOR
CC FIBRINOGEN, FACTOR X AND ICAM1. IT RECOGNIZES P1 AND P2 PEPTIDES
CC OF FIBRINOGEN GAMMA CHAIN.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-M
CC ASSOCIATES WITH BETA-2.
CC -!- SUBCELLULAR LOCATION: PREDOMINANTLY EXPRESSED IN MONOCYTES AND
CC GRANULOCYTES.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD11b entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd11b.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

Query Match 99.1%; Score 5823.5; DB 1; Length 1152;
 Best Local Similarity 98.8%; Pred. No. 0;
 Matches 1123; Conservative 8; Mismatches 5; Indels 1; Gaps 1;

QY 1 FNLDTENAMTFOENARGFQSQVVOGGSRVVGAPQEIIVANQKSLVQCDYSTGSCPEI 60
 DB 17 FNLDTENAMTFOENARGFQSQVVOGGSRVVGAPQEIIVANQKSLVQCDYSTGSCPEI 76

QY 61 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHTCSENTVYKGLCLFGLSGLNQKQK 120
 DB 77 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHTCSENTVYKGLCLFGLSGLNQKQK 136

QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKELVSTIMBQLKSKTLFSLMYSEEF 180
 DB 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKELVSTIMBQLKSKTLFSLMYSEEF 196

QY 181 RTHFTKPEKONNPNRSLKPTTOLLGRTHATGLRKVRELFNITNGARKNAKILFL 240
 DB 197 RTHFTKPEKONNPNRSLKPTTOLLGRTHATGLRKVRELFNITNGARKNAKILVYI 256

QY 241 TDGEKFGDPLGYEDVPEIDREGVIRYVLGFDAPRSEKSRQELNAVSKPPRHDVFOAN 300
 DB 257 TDGEKFGDPLGYEDVPEIDREGVIRYVLGFDAPRSEKSRQELNAVSKPPRHDVFOAN 316

QY 301 NFEALKTQVQOLREKIFAETGCTGTSSSFHEMSQSGESAITSNGELLSTVGSYDAG 360
 DB 317 NFEALKTQVQOLREKIFAETGCTGTSSSFHEMSQSGESAITSNGELLSTVGSYDAG 376

QY 361 GYFLYTSKEKSFINMTRVDSMDNDAYLGAAAILLNRVQSLVGLGAPYQIHGLVAMER 420
 DB 377 GYFLYTSKEKSFINMTRVDSMDNDAYLGAAAILLNRVQSLVGLGAPYQIHGLVAMER 436

QY 421 QNTGHWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIAGPHYETQRGQVSVCP 480
 DB 437 QNTGHWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIAGPHYETQRGQVSVCP 496

QY 481 PRGORARWQCDVLYGECQGPGRFGAALTGLDVNGDKLTDLVAGPDEEDNRGAVILF 540
 DB 497 PRG-RARWQCDVLYGECQGPGRFGAALTGLDVNGDKLTDLVAGPDEEDNRGAVILF 555

QY 541 HGTSGSGISPSHRSQRIAGSKLSPRLQYFQSGGQDLTMDGLVLTGGAQGHVLLLRSQ 600
 DB 556 HGTSGSGISPSHRSQRIAGSKLSPRLQYFQSGGQDLTMDGLVLTGGAQGHVLLLRSQ 615

QY 601 PVLRVKAIEMFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLRGQIQSVVT 660
 DB 616 PVLRVKAIEMFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLRGQIQSVVT 675

QY 661 YDLALDSGRPHSAVNETKNSRRTQVLGTLQTCETLKLQLPNCIEDPVSPIVLRNF 720
 DB 676 YDLALDSGRPHSAVNETKNSRRTQVLGTLQTCETLKLQLPNCIEDPVSPIVLRNF 735

QY 721 SLVGTPLSAFGLNRPVLAEDARLFTALPPFEKNCQNDNICQDDLSITPSFMSLCLVVG 780
 DB 736 SLVGTPLSAFGLNRPVLAEDARLFTALPPFEKNCQNDNICQDDLSITPSFMSLCLVVG 795

QY 781 GRPEFNVTVVRNDGDSYRTQVTFPPPLDLSYRKVSTLQNSORSQSWELACESASSTEV 840
 DB 796 GRPEFNVTVVRNDGDSYRTQVTFPPPLDLSYRKVSTLQNSORSQSWELACESASSTEV 855

QY 841 SGALKSTCSINHIPIFENSEVFNITFDVDSKASLGKLLKANKVTSNNMPTNKTEF 900
 DB 856 SGALKSTCSINHIPIFENSEVFNITFDVDSKASLGKLLKANKVTSNNMPTNKTEF 915

QY 901 QLELPVKYAVVWVTSHGVSQKYNLTASNTSRVMQHOYQVSNLQORSLSPLSLFLVLPV 960
 DB 916 QLELPVKYAVVWVTSHGVSQKYNLTASNTSRVMQHOYQVSNLQORSLSPLSLFLVLPV 975

QY 961 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCSTAVCORIQCDIP 1020
 DB 976 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCSTAVCORIQCDIP 1035

RESULT 2

ITAM_MOUSE

ID ITAM_MOUSE STANDARD; PRT; 1153 AA.

AC P0555;

DT 01-NOV-1988 (Rel. 09, Created)

DT 01-FEB-1991 (Rel. 17, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha

subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor MO1).

GN ITGAM.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN (1)

RP SEQUENCE FROM N.A.

RX MEDLINE=89312584; PubMed=3044779;

RA Pyela R.;

RT "Amino acid sequence of the murine Mac-1 alpha chain reveals homology

with the integrin family and an additional domain related to von

Willebrand factor.";

RL EMBO J. 7:1371-1378 (1988).

RN (2)

RP SEQUENCE OF 11-45 FROM N.A.

RC STRAIN=BALE/c; TISSUE=Spleen;

RX MEDLINE=86287312; PubMed=2942940;

RA Sastre L., Roman J.M., Teplow D.B., Dreyer W.J., Gee C.E.,

Larson R.S., Roberts T.M., Springer T.A.;

"A partial genomic DNA clone for the alpha subunit of the mouse

complement receptor type 3 and cellular adhesion molecule Mac-1.";

Proc. Natl. Acad. Sci. U.S.A. 83:5644-5648 (1986).

RN (3)

RP SEQUENCE OF 17-28.

RX MEDLINE=85188276; PubMed=3887182;

RA Springer T.A., Teplow D.B., Dreyer W.J.;

"Sequence homology of the LFA-1 and Mac-1 leukocyte adhesion

glycoproteins and unexpected relation to leukocyte interferon.";

Nature 314:540-542 (1985).

CC -1- FUNCTION: INTEGRIN ALPHA-M/BETA-2 IS IMPLICATED IN VARIOUS

ADHESIVE INTERACTIONS OF MONOCYTES, MACROPHAGES AND GRANULOCYTES

AS WELL AS IN MEDIATING THE UPTAKE OF COMPLEMENT-COATED PARTICLES.

IT IS IDENTICAL WITH CR-3, THE RECEPTOR FOR THE IC3B FRAGMENT OF

THE THIRD COMPLEMENT COMPONENT. IT PROBABLY RECOGNIZES THE R-G-D

PEPTIDE IN C3B. INTEGRIN ALPHA-M/BETA-2 IS ALSO A RECEPTOR FOR

FIBRINOGEN, FACTOR X AND ICAM1. IT RECOGNIZES P1 AND P2 PEPTIDES

OF FIBRINOGEN GAMMA CHAIN. ALPHA-M/BETA-2 PLAY A CRITICAL ROLE IN

MAST CELL DEVELOPMENT AND IN IMMUNE COMPLEX-MEDIATED

GLOMERULONEPHRITIS. MICE EXPRESSING A NULL MUTATION OF THE ALPHA-M

SUBUNIT GENE DEMONSTRATE INCREASE IN NEUTROPHIL ACCUMULATION, IN

RESPONSE TO A IMPAIRED DEGRADATION AND PHAGOCYTOSIS, EVENTS THAT

APPARENTLY ACCELERATE APOPTOSIS IN NEUTROPHILS. THESE MICE DEVELOP

OBESITY.

CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN. ALPHA-M

ASSOCIATES WITH BETA-2.

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND

GRANULOCYTES.

CC -1- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS

WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.

CC -1- SIMILARITY: Belongs to the integrin alpha chain family.

CC -1- SIMILARITY: Contains 1 VWFA domain.

CC -1- SIMILARITY: Contains 7 FG-GAP repeats.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC -----
DR ENBL; X07640; CAA30479.1; -.
DR EMBL; M14293; AAA39484.1; -.
DR PIR; S00551; S00551.
DR HSP; P11215; IABX.
DR XGD; MGI:96607; Itgam.
DR InterPro; IPR00413; Integrin_alpha.
DR InterPro; IPR02035; VWF_A.
DR Pfam; PF01839; FG-GAP 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWA; 1.
DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Signal; Calcium; Repeat.
FT SIGNAL 1 16
FT CHAIN 17 1155 INTEGRIN ALPHA-M.
FT DOMAIN 17 1105 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1106 1129 POTENTIAL.
FT DOMAIN 1130 1153 CYTOPLASMIC (POTENTIAL).
FT REPEAT 31 84 FG-GAP 1.
FT REPEAT ? 7 FG-GAP 2.
FT DOMAIN 164 350 VWA.
FT REPEAT 337 400 FG-GAP 3.
FT REPEAT 401 452 FG-GAP 4.
FT REPEAT 454 515 FG-GAP 5.
FT REPEAT 517 575 FG-GAP 6.
FT REPEAT 580 632 FG-GAP 7.
FT CA_BIND 465 473 POTENTIAL.
FT CA_BIND 529 537 POTENTIAL.
FT CA_BIND 592 600 POTENTIAL.
FT SITE 1132 1136 GPFKR MOTIF.
FT DISULFID 66 73 BY SIMILARITY.
FT DISULFID 105 123 BY SIMILARITY.
FT DISULFID 654 711 BY SIMILARITY.
FT DISULFID 770 776 BY SIMILARITY.
FT DISULFID 999 1023 BY SIMILARITY.
FT DISULFID 1028 1033 BY SIMILARITY.
FT CARBOHYD 58 58 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 86 86 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 391 391 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 696 696 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 734 734 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 772 772 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 801 801 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 881 881 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 907 907 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 941 941 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 980 980 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 994 994 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1022 1022 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1045 1045 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1051 1051 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1076 1076 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1153 AA; 127480 MW; 178DB989AECB0343 CRC64;

Query Match 75.6%; Score 4447; DB 1; Length 1153;
Best Local Similarity 73.5%; Pred. No. 1.5e-292;
Matches 836; Conservative 147; Mismatches 153; Indels 2; Gaps 2;

QY 1 FNLDTENAMTPOENAEFGQSVVQLGSGRVVVCAGPOEIVAAQNRSLYQCDYSTGSCBPI 60
DB 17 FNLDTEHPMTPOENAEFGQSVVQLGSGRVVVCAGPOEIVAAQNRSLYQCDYSTGSCBPI 76

```

```

QY 61 RLOVPEAVNMSLGLSLAATTSPOLLACGPTVHOTCSENTVVKGLCLFPGSLRQPOK 120
DB 77 PLOVPEAVNMSLGLSLAVTFPOLLACGPTVHONCKENTYVNGLCVLFPGSLRPPQQ 136
QY 121 PPEALRGCPQSDSDIAFLIDSGSII PHDPRMKELVSTIMEOLKSKXTLSLMQYSBFF 180
DB 137 PPEALRGCPQSDSDIVFLIDSGSINNIDFQMKEFVSTVMEQFKSKTSLFSLMQYSBFF 196
QY 181 RIHFTFEKFFONNPNRSLIKPTTOLLGRTHATGLRKVVVRLENTNIGARKNAKIFLL 240
DB 197 RIHFTFENFKNPSRSHVSPDKQLNGRTKTASGIRKVVVRLEFKTNGARENAAILVVI 256
QY 241 TDGEKFGDPLGYEDVIPELDREGVIRYVLGPDAPRSEKSRQELHNTVASKPRDHFVQAN 300
DB 257 TDGEKFGDPLDYKDVIPEDRAGVIRYVIGVGNFKNPKQSRRELDTIASKPAGHVFQVD 316
QY 301 NPEALKTVQNLREKI PAIEGTQTGSSSFEHEMSQEGFSAAITNSGPLLSTVGSVDWAG 360
DB 317 NPEALNTIQNLQEKI PAIEGTQTGSSSFEHEMSQEGFSAAITNSGPLLSTVGSVDWAG 376
QY 361 GVFLYTSXEKSTFINMTRVSDMNDAYLGYAAAI LRNRVQSLVGLGAPRYQHIGLVAMFR 420
DB 377 GAFLYTSXDKVTFINMTRVSDMNDAYLGYASAVILNRNRVQSLVGLGAPRYQHIGLVAMFR 436
QY 421 QNTGHWESNANVKTQICAYEGASILCSVDVDSNSTDLVILGAPHYEOTRGQGVSCPL 480
DB 437 ENFGTWEPHTSIKSGISQISYFGASILCSVDVDSNSTDLVILGAPHYEOTRGQGVSCPL 496
QY 481 PRGQARWQCDVILYGEQGPWGRFGAALT VLGVDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
DB 497 PRG-RARWQCEALLHGDQHPAGHFGAALT VLGVDVNGDKLTDVAIGAPGEENQAVYIF 555
QY 541 HGTSGSGLSPSHSRIAGSKLSPRIQYFGQSLGSGQDLTMDGLVDLT VGAQHVLRLRSQ 600
DB 556 YGASTASLSASHSRITGAHFSQGLQYFGQSLGSGKDLTMDGLMDL VAGAQHVLRLRSQ 615
QY 601 PVLRVKAIEMFNPREVARNVPCNDQVVKGEAGVRVCLHVQKSTRDLREGQIQSVVT 660
DB 616 PVLRLKATWEPSPKKVARSVPACQEVLLKNKDAEVRVCLVRKNTKDLREGDIQSVVT 675
QY 661 YDLALDSGRPHSRVFNETKSTRTQTOVLGLTCTETKLQPCNIEDDPSPVILRLNF 720
DB 676 YDLALDPVRSIRAFDFETKNTTRTQTOVLGLMOKCETLKLILPDCVDSVSPILRLNY 735
QY 721 SLVGTPLSAPGNLRPVLAEDAQRLFTALFPPEKKGNDNICODDISITFSFMSDCLVVG 780
DB 736 TLVGEPLSRGNLRPVLAEDAQRLFTALFPPEKKGNDNICODDISITFSFMSDCLVVG 795
QY 781 GPRFNVTVVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRL-ACRSASSTE 839
DB 796 GPQDFNMSVTLRNDGDSYGTQVTVYTPSGLSYRKDSASQNP LTKPWFVKPABSSSSSE 855
QY 840 VSGALKSTSCSINHPI PPESEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKT 899
DB 856 GHGALKSTTWNINHPI PPNSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKT 915
QY 900 FQLELPVKYAVIMVTSHGVSTKYLNFTASENTRVMOHQYQVSNLQORSPLI SILVFLVP 959
DB 916 FQLELPVKYAVIMVTSHGVSTKYLNFTASENTRVMOHQYQVSNLQORSPLI SILVFLVP 975
QY 960 VRLNQVTIWDPEPOTFSENSTCHTKERLPSHSDLAELRKA PVNCSIAVCORICDI 1019
DB 976 VQINNVTVWDHPQVIFSNLSAACHTQKSPHNFDRDLERTPVLCNSVAVCKRIQCDL 1035
QY 1020 PFFGIQBEFNATLKNLSFDWYIKTSHNHLIVSTABILFNDVSTFTLLPGQCAFVRSOTE 1079
DB 1036 PSFNTQBIENVTLKNLSFDWYIKTSHNHLIVSTABILFNDVSTFTLLPGQCAFVRSOTE 1095
QY 1080 TKVEPPEVNPPLPLVGVSSVGHLLALITAAALKGPFKQYKDMSEGGCPGPAEPQ 1137
DB 1096 TKVPEVYVNPVPLVGVSSVGHLLALITAGLYKLGPFKQYKDMSEGGCPGPAEPQ 1153

```


Db 138 LPVSRQCPROEQDIVFLIDSGSISRNFMATMNFVRAVISQPRSTQFSLQFSNKF 197
Qy 181 RIHTFKEFQNNPNRSLRIKFIITQGLRTHITATGLRVKVRLEFNTMGARKNAFKILFL 240
Db 198 QTHFTFEFRSTNPLSILASVHLOQFTVATATQNVVRLPHASVGARRDATKILIVI 257
Qy 241 TDGKFGDPLGYEDVIBELDEGVIRVVLGDAFRSEKROELNTVASPPDRHVPQAN 300
Db 258 TDGKFGDPLGYEDVIBELDEGVIRVVLGDAFRSEKROELNTVASPPDRHVPQAN 317
Qy 301 NFEALTKVQNLREKIFAIBGTQTGSSSSFEHMSQGFSAATISNGPILLSVGSVDWAG 360
Db 318 DFDALKIQNLREKIFAIBGTQTGSSSSFEHMSQGFSAATISNGPILLSVGSVDWAG 377
Qy 361 GVPLYTSEKSTPNMTVDSDMDAYLGYAAAILLRVQSLVGLGAPRYOHILGVAMER 420
Db 378 GAFLLPNNMSTPFINMSQENVNDSYGLSTELALWGVQSLVGLGAPRYOHTGKAVIF 437
Qy 421 QNTGMESNANVKTQIGAFYGASLCSVDVDSNGSTDVLIGAPHYVYQTRGGQVSCPL 480
Db 438 QVSRQWRKAEVTCQIGSYFGSLCSVDVDSNGSTDVLIGAPHYVYQTRGGQVSCPL 497
Qy 481 PRQARWQCDALYEGQGFWRGFGAALTIVLGVNGDKLTDVAIGAPGEDNRGAVYLF 540
Db 498 PRQARWQCDALYEGQGFWRGFGAALTIVLGVNGDKLTDVAIGAPGEDNRGAVYLF 556
Qy 541 HGTSGSISFHSQRIAGSKLSPLOYFGOSLSGGDLTMDGLVDTLVGAGHVLILRSQ 600
Db 557 HGVLFPSISFHSQRIAGSKLSPLOYFGOSLSGGDLTMDGLVDTLVGAGHVLILRSQ 616
Qy 601 PVLRVKAEMFNPREVARNVFCNDQVYKGEAGEVRVCLHVQKSTRDLRREGQIQSVTF 660
Db 617 PVLRVKAEMFNPREVARNVFCNDQVYKGEAGEVRVCLHVQKSTRDLRREGQIQSVTF 676
Qy 661 YDLALDSGRPHSAVENETKNSRTOVLGLTQCTETKLPNCIEQVPSVILRLNF 720
Db 677 LDALDPERLSPRATQETKNSRVRVLGKAHCNENLLPSCVEDSVPTITURLNF 736
Qy 721 SLVGTPLSAFNGRLPRLAEDAQRLTALFPFERNCGNDNICQDLSITFSFMSLDCLVWG 780
Db 737 TLVCKPLLAFLNRLPMLAQAQYFASLPPFKNCGADHICQDNLGSPFGLKSLVWG 796
Qy 781 GPREFNVTVVRNDGDSYRTQVTFPPDLDSVRKYSTUQONQBSQSWLACASSTEV 840
Db 797 SNLELNAEVMVNDGDSYRTQVTFPPDLDSVRKYSTUQONQBSQSWLACASSTEV 854
Qy 841 SGALKTSICSNIPFIPENSEVFNITFDVDSKASLGNKLLKANVTSENNPRTNKTFF 900
Db 855 SGTWSTSCRNELIFRGGAQITFLATFVDSPKAVLGRLLLTANVSENNPTFSKITFF 914
Qy 901 QLELPVXYAYVYVTSHEVSTKYLNTAS-ENTSRVMOHQYQVSNLQGRSLPISLVFLVP 959
Db 915 QLELPVXYAYVYVTSHEVSTKYLNTAS-ENTSRVMOHQYQVSNLQGRSLPISLVFLVP 974
Qy 960 VRLNQTVIMDRPOVTFSEMLSSSTCHTKERLPSSHDFLAEARAPVNVNCIAVCORICDI 1019
Db 975 VELNQAEVMDVESHVQNPSPSCSEKTAAPASDFLAHQKPNVLDCSAGCLAFRCVD 1034
Qy 1020 PFGIOBEFNATKGNLSFDWYIKTSHNLLIVSTAEILFNDSVFTLLPQGAQVRSQTE 1079
Db 1035 PFSVOEELDTLKGNSLPCWVRQILQKXVSVVVAEITFDTSVSOLPQGAQVRSQTE 1094
Qy 1080 TKVEPEVFNPLIVGSSVGLLALITAAALYKLGPFKFOYKDMWSE 1128
Db 1095 TVLEKVKVNPPLIVGSSVGLLALITAAALYKLGPFKFOYKDMWSE 1143

RESULT 4

ID ITAD_HUMAN
AC Q13349; Q15575; Q15576;
BT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-D precursor (Leukointegrin alpha D) (CD11d) (ADB2).
GN ITGAD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RX MEDLINE=96111956; PubMed=8777714;
RA Van der Vieren M., Le Trong H., Wood C.L., Moore P.F., St John T.,
RA Staunton D.E., Gallatin W.M.;
RT "A novel leukointegrin, alpha d beta 2, binds preferentially to ICAM-
RT 3.";
RL Immunity 3:683-690 (1995).
RN [2]
RP SEQUENCE OF 1-235 FROM N.A.
RX MEDLINE=20187620; PubMed=10722744;
RA Noti J.D., Johnson A.K., Dillon J.D.;
RT "Structural and functional characterization of the leukocyte integrin
RT gene CD11d. Essential role of Sp1 and Sp3.";
RL J. Biol. Chem. 275:8959-8969 (2000).
RN [3]
RP SEQUENCE OF 457-537; 571-602; 633-666; 788-834 AND 910-1125 FROM N.A.
RX MEDLINE=96257236; PubMed=8666289;
RA Wong D.A., Davis E.M., LeBeau M., Springer T.A.;
RT "Cloning and chromosomal localization of a novel gene encoding a human
RT beta 2-integrin alpha subunit.";
RL Gene 171:291-294 (1996).
RN [4]
RP INTERACTION WITH VCAM1.
RX MEDLINE=99059842; PubMed=9841932;
RA Grayson M.H., Van der Vieren M., Sterbinsky S.A., Michael Gallatin W.,
RA Hoffman P.A., Staunton D.E., Sochner B.S.;
RT "alpha2beta2 integrin is expressed on human eosinophils and functions
RT as an alternative ligand for vascular cell adhesion molecule 1
RT (VCAM-1).";
RL J. Exp. Med. 188:2187-2191 (1998).
RN [5]
RP INTERACTION WITH VCAM1.
RX MEDLINE=99370002; PubMed=10438935;
RA Van der Vieren M., Crowe D.F., Hoekstra D., Vazeux R., Hoffman P.A.,
RA Grayson M.H., Sochner B.S., Gallatin W.M., Staunton D.E.;
RT "The leukocyte integrin alpha D beta 2 binds VCAM-1: evidence for a
RT binding interface between I domain and VCAM-1.";
RL J. Immunol. 163:1984-1990 (1999).
CC [1]- FUNCTION: INTEGRIN ALPHA-D/BETA-2 IS A RECEPTOR FOR ICAM3 AND
CC VCAM1. MAY PLAY A ROLE IN THE ATHEROSCLEROTIC PROCESS SUCH AS
CC CLEARING LIPOPROTEINS FROM PLAQUES AND IN PHAGOCYTOSIS OF BLOOD-
CC BORNE PATHOGENS, PARTICULATE MATTER, AND SENESCENT ERYTHROCYTES
CC FROM THE BLOOD.
CC [2]- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-D
CC ASSOCIATES WITH BETA-2.
CC [3]- SUBCELLULAR LOCATION: Type I membrane protein.
CC [4]- TISSUE SPECIFICITY: EXPRESSED MODERATELY ON MYELOMONOCYTIC CELL
CC LINES AND SUBSETS OF PERIPHERAL BLOOD LEUKOCYTES AND STRONGLY ON
CC TISSUE-SPECIALIZED CELLS, INCLUDING MACROPHAGES FOAM CELLS WITHIN
CC ATHEROSCLEROTIC PLAQUES, AND ON SPLENIC RED PULP MACROPHAGES.
CC [5]- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWF A DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC [6]- SIMILARITY: Belongs to the integrin alpha chain family.
CC [7]- SIMILARITY: Contains 1 VWF A domain.
CC [8]- SIMILARITY: Contains 7 FG-GAP repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)
CC or send an email to license@isb-sib.ch.
CC -----

DR	EMBL; U37028; AAB38547.1; -	1	17	POTENTIAL.
DR	EMBL; U40274; AAB60634.1; -	18	1162	INTEGRIN ALPHA-D.
DR	EMBL; U40275; AAB60635.1; -	18	1100	EXTRACELLULAR (POTENTIAL).
DR	EMBL; U40276; AAB60636.1; -	1101	1124	POTENTIAL.
DR	EMBL; U40277; AAB60637.1; -	1125	1162	CYTOSOLASMIC (POTENTIAL).
DR	EMBL; U40278; AAB60638.1; -	32	85	FG-GAP 1.
DR	EMBL; U40279; AAB60639.1; -	?	?	FG-GAP 2.
DR	EMBL; U40278; AAB60638.1; JOINED.	150	332	WFA.
DR	EMBL; AF187881; AAF62875.1; -	350	400	FG-GAP 3.
DR	HSSP; PL1215; 1A8X	401	452	FG-GAP 4.
DR	Genew; HGNC:6146; ITGAD.	454	516	FG-GAP 5.
DR	MIM; 602453; -	518	576	FG-GAP 6.
DR	GO; GO:0006305; -	581	633	FG-GAP 7.
DR	GO; GO:0004895; F:cell adhesion receptor activity; TAS.	581	633	POTENTIAL.
DR	GO; GO:0016337; P:cell-cell adhesion; NAS.	465	473	POTENTIAL.
DR	GO; GO:0007160; P:cell-matrix adhesion; NAS.	530	538	POTENTIAL.
DR	GO; GO:0006955; P:immune response; NAS.	593	601	POTENTIAL.
DR	InterPro; IPR000413; Integrin_alpha.	1127	1131	GFGR MOTIF.
DR	InterPro; IPR002035; VWF_A.	67	74	BY SIMILARITY.
DR	Pfam; PF01839; FG-GAP_3.	106	124	BY SIMILARITY.
DR	Pfam; PF00357; integrin_A_1.	655	710	BY SIMILARITY.
DR	Pfam; PF00092; vwa; 1.	769	775	BY SIMILARITY.
DR	PRINTS; PRO1185; INTEGRINA.	846	861	BY SIMILARITY.
DR	PRINTS; PRO0453; VWFADOMAIN.	994	1018	BY SIMILARITY.
DR	SMART; SM00191; Int_alpha; 4.	1023	1028	BY SIMILARITY.
DR	SMART; SM00327; Vwa; 1.	59	59	N-LINKED (GLCNAC. -) (POTENTIAL).
DR	PROSITE; PS00242; INTEGRIN_ALPHA; 1.	87	97	N-LINKED (GLCNAC. -) (POTENTIAL).
DR	PROSITE; PS00234; VWFA; 1.	99	99	N-LINKED (GLCNAC. -) (POTENTIAL).
DR	Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;	391	391	N-LINKED (GLCNAC. -) (POTENTIAL).
DR	Signal; Repeat; Calcium;	691	691	N-LINKED (GLCNAC. -) (POTENTIAL).
DR	Magnesium.	733	733	N-LINKED (GLCNAC. -) (POTENTIAL).
DR	SIGNAL	873	873	N-LINKED (GLCNAC. -) (POTENTIAL).
DR	CHAIN	957	957	N-LINKED (GLCNAC. -) (POTENTIAL).
DR	DOMAIN	1046	1046	N-LINKED (GLCNAC. -) (POTENTIAL).
DR	TRANSMEM	500	500	MISSING (IN REF. 2).
DR	DOMAIN	515	518	GHWP -> ATP (IN REF. 2).
DR	REPEAT	825	825	L -> V (IN REF. 2).
DR	REPEAT	984	984	V -> A (IN REF. 2).
DR	REPEAT	1162	AA; 126885 MW; F296AIA35455077D CRC64;	


```

FT CARBOHYD 65 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 89 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 188 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 649 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 670 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 726 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 730 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 862 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 885 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 897 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 1060 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 1071 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT VARSPLIC 954 Q -> QGVRLGVEMQTSKQILCEPAGDAHTVGAQGEPLCP
PWGSEAPRDNIRAGPCR (in isoform 2).
/FTIGVSP 002738.
R -> W (IN REF. 1 AND 2).
Y -> I (IN REF. 2).

FT CONFLICT 214
FT CONFLICT 660
FT STRAND 155
FT STRAND 164
FT TURN 165
FT TURN 166
FT HELIX 169
FT TURN 186
FT STRAND 191
FT STRAND 202
FT HELIX 208
FT HELIX 217
FT TURN 222
FT STRAND 229
FT HELIX 233
FT TURN 244
FT HELIX 247
FT TURN 250
FT TURN 253
FT STRAND 256
FT HELIX 274
FT TURN 277
FT STRAND 280
FT HELIX 288
FT TURN 293
FT TURN 299
FT HELIX 300
FT HELIX 307
FT TURN 311
FT TURN 318
FT HELIX 319
FT TURN 329
SQ SEQUENCE 1170 AA; 128819 MW; 39A7AF92EF286FC0 CRC64;

Query Match 26.1%; Score 1536.5; DB 1; Length 1170;
Best Local Similarity 34.2%; Pred. No. 1.2e-95;
Matches 401; Conservative 209; Mismatches 463; Indels 101; Gaps 37;

QY 1 FNLDTENAMTFQ--ENARGFGQSVQVQLQGSRRVVVGAPOEIVAAHQSGSLYQCDSYSGSCE 58
DB 26 YNLDVRGARSFSPFRAGRHFGYEVLVQV-GNGVIVGAPGE---GNSGTSLYQCDSYSGHCL 81
QY 59 PTLQVPEAVNMSGLSLAATSPOLLACGGTVTQTCSENVTYVGLCLPLFGSNLR--- 115
DB 82 PVTLR-GSNYTSKYLGMILATDPTDGSILACDPLGSRCTDQNTYLSGLCYLFRQNLQGP 140
QY 116 -QQPKPFPALRGCPQEDSDIAFLIDGSGSIIIPDFRRMKELVSTIMEOLKSKSLFSLM 174
DB 141 LQRPQFPQECIKG----NVDLVFLFGSVLSQDFQKILDFPKDVKKLSNYSYQPAAY 196
QY 175 QYSEFRHIFTKPEFQNNPNRSLIKPITOLLGRHTATGLRKVRVLEFNITNGAKKNAP 234
DB 197 QFSTYSKTEPFDSDYVKKRQDPDALLKHVKHMLLTNTFGAINTVATEVREBELGAPDAT 256
QY 235 KILPLTDEGKFDPLGYEDVIELOREGVIRVILGFDAPFSEKSRQSLNTVASKPPED 294
DB 257 KVLIIITDGS--ATDSGNIDAAND-----IIRVIIGIKGHFQTKESQETILKFPASKPASE 309
QY 295 HVFQANNFEALTKTVQNLREKIPAEIGTQTGSSSSFEHEMSQEGFSAATISNGPLISTV 354

```

```

DB 310 FVKILDTFEKLDLFTLQKKIYVIEGTSKODLTSPNMLSSSGISADLSRGHVVGVAG 369
QY 355 SYDNAGGVF-LYTSKEKSTFINMTVDSDMDAYLGYAAA-IILNRVQSLVLGAPRYOH 412
DB 370 AXDNAGGFLDLKADLQDDTFIGNEPITFEVRAGLYGVTVWLPSPQKTSLLASGAPRYOH 429
QY 413 IGLVAMER--QNTGMWESNANVKQIQIGAYGASICSVVDVDSNGSTDLVLGAPHYEOT 470
DB 430 MGRVLLFQEPQGGGHWSQVTHGTQIGSYFGELCGVDVDQDGETELLIGAPLYFGQ 489
QY 471 RGGQVSVCPPLPGQPARWQCDAY--LYGEOQOPNCFGAALTVLGDVNGDKLTDLVAIGAP 528
DB 480 RGRVFIY-----QRRQLGFBEVSELOQDPGLRFGAITALDINDGLVDVAVGAP 544
QY 529 GEDNRGAVYLFPHGTSGSGISFSGRIAGSKLSPLQYFGQSLSGGODLTMDGLVDLV 588
DB 545 LEE--QGAVYIFNGRHG-GLSPQSPRSORIEGTQVLSGIQWFGRSIHGVKDLGDLAVAV 601
QY 589 GAGCHVLLLRQPVLRVKAINMEFNPREVARNVFECDQV-KGKEAGEVRVCLHVOKSTR 647
DB 602 GABSQIVLSRPVVDVMTLMSFSPAEPVHEVECSYSTSNRMEGVNITICFQI-KSLY 660
QY 648 DRLREGQIQSVVTVYDLALDSGRPHSRVAFNFTKSTRQTQVTLGLTQTCETLKLQLENCI 707
DB 661 PQF-QGRLVANLTYTLQDGHTRRGLFPGGRHLERNAIVT-TSMSCDTPSPFPVVCV 718
QY 708 EDPVSPVILRLNFSI---VGTPLS--AFGN-----LRPVLAEDAQRLTALPPEKNCN 757
DB 719 QDLISFINVSLNFSLWEEEGTPRQRAQKQDIPPIRLRSLHSETWEI-----PFEKNCE 773
QY 758 DNICODDLSITFSFMSLDCLVVGGRPFNVTWVNDGEDSYRTQVTPFFPDLDSVRKVS 817
DB 774 DKCEANLRVSPSPARSALALTAPASLSVELSLNLEEDAYVQLDLHFPGLSPKVE 833
QY 818 TLQNRQSRWRLACES--ASSTEVGALKTSCSINHIPIIPENSEVTNITFDVDSKAS 875
DB 834 ML---KPHSQIPVSCCELPEESRLSRAL---SCNVSSPIFKAGHSVALQMMENFVANS 887
QY 876 LGNKLLKXANTVSENN---MPTNKTEFQLELPVKYAVVMVVTSGHVSXTYLNFTASEN 931
DB 888 WDSVVELHANTCNEDSDLEDNSATTI---IPILYPINILIQOEDSTLYVSTPKGP 944
QY 932 TSVVMQHOYQV---SNLQGRSLP-ISLVFLVPLRLNQTIVMDRPQVTFSENLSSTCHTK- 986
DB 945 KIHQVKHMYQVRIQPSIHDHNIPTLEAVVGVQPPESEGPITHQWSVQVMEPPV--PCHVED 1002
QY 987 -ERLPSHSD--FLAELRKAPVUNCIAVCQRIQCDIPFGIOBEBFNATLKNLSFDWYIK 1043
DB 1003 LERLPDAAEPCLPGALFPCPVV-----FRQELVQVIGLELVGEIE 1044
QY 1044 TSHNELLIVSTABILFNDVSFTLLPGQGAFFRSQTFETKVPFEVENPLPLIVGSSVGGLL 1103
DB 1045 AS-SMFSLCSSLSISFNSSKHFLYGSNASL-AQVVMKVDVYVYKQMLYLYVLSGIGELL 1102
QY 1104 LLALITAAALYKLGPKKQYKDMSEG-CPRGAP 1136
DB 1103 LLLLIPTVLYKVGFPKRLKEMEAGRPNGIP 1136

RESULT 6
ITAL MOUSE
ID ITAL MOUSE STANDARD; PRT: 1163 AA.
AC P24063;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-L precursor (Leukocyte adhesion glycoprotein LFA-1 alpha chain) (Leukocyte function associated molecule 1, alpha chain) (CD11a).
DE ITGAL OR LFA-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91269576; PubMed=2051027;
RA Kaufmann Y., Tseng B., Springer T.A.;
RT "Cloning of the murine lymphocyte function-associated molecule-1
alpha-subunit and its expression in COS cells.";
RL J. Immunol. 147:369-374(1991).
RN [2]
RP SEQUENCE OF 24-42.
RX MEDLINE=9518276; PubMed=3887182;
RA Springer T.A., Teplow D.B., Dreyer W.J.;
RT "Sequence homology of the LFA-1 and Mac-1 leukocyte adhesion
glycoproteins and unexpected relation to leukocyte interferon.";
RL Nature 314:540-542(1985).
CC -!- FUNCTION: INTEGRIN ALPHA-L/BETA-2 IS A RECEPTOR FOR ICAM1, ICAM2,
ICAM3 AND ICAM4. IS INVOLVED IN A VARIETY OF IMMUNE PHENOMENA
INCLUDING LEUKOCYTE-ENDOTHELIAL CELL INTERACTION, CYTOTOXIC T-CELL
MEDIATED KILLING, AND ANTIBODY DEPENDENT KILLING BY GRANULOCYTES
AND MONOCYTES. NICE EXPRESSING A NULL MUTATION OF THE ALPHA-L
SUBUNIT GENE DEMONSTRATE IMPAIRED TUMOR REJECTION AND IMPAIRED
LEUKOCYTES RECRUITMENT.
CC -!- SUBUNIT: HETEROIDIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-L
ASSOCIATES WITH BETA-2.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: LEUKOCYTES.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSET) IS A VWFA DOMAIN. INTEGRINS
WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M60778; AAA39426.1; --
DR PIR: I56126; I56126.
DR HSSP: P20701; 1LFA.
DR MGD: MGI:96606; Itgal.
DR InterPro: IPR000413; Integrin_alpha.
DR InterPro: IPR002035; VWFA.
DR Pfam: PF01839; FG-GAP; 3.
DR Pfam: PF00357; Integrin_A; 1.
DR Pfam: PF00092; vwa; 1.
DR PRINTS: PR01185; INTEGRINA.
DR PRINTS: PR00453; VWFADOMAIN.
DR SMART: SM00191; Int_alpha; 5.
DR SMART: SM00327; VWFA; 1.
DR PROSITE: PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE: PS0234; VWFA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Signal; Magnesium; Calcium;
KW Repeat.
FT SIGNAL 1 23
FT CHAIN 24 1163
FT DOMAIN 24 1084
FT TRANSMEM 1085 1108
FT DOMAIN 1109 1163
FT REPEAT 39 88
FT REPEAT ? ?
FT DOMAIN 148 334
FT REPEAT ? ?
FT REPEAT 399 454
FT REPEAT 455 514
FT REPEAT 516 573
FT REPEAT 576 628
FT CA_BIND 466 474
FT SIGNAL 1 23
FT CHAIN 24 1163
FT DOMAIN 24 1084
FT TRANSMEM 1085 1108
FT DOMAIN 1109 1163
FT REPEAT 39 88
FT REPEAT ? ?
FT DOMAIN 148 334
FT REPEAT ? ?
FT REPEAT 399 454
FT REPEAT 455 514
FT REPEAT 516 573
FT REPEAT 576 628
FT CA_BIND 466 474

FT	CA_BIND	528	536	POTENTIAL.
FT	CA_BIND	588	596	POTENTIAL.
FT	SITE	1111	1115	GFPR MOTIF.
FT	DISULFID	70	77	BY SIMILARITY.
FT	DISULFID	103	126	BY SIMILARITY.
FT	DISULFID	147	139	BY SIMILARITY.
FT	DISULFID	651	705	BY SIMILARITY.
FT	DISULFID	767	773	BY SIMILARITY.
FT	DISULFID	840	856	BY SIMILARITY.
FT	DISULFID	993	1009	BY SIMILARITY.
FT	DISULFID	1017	1048	BY SIMILARITY.
FT	CARBOHYD	85	86	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	185	185	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	270	270	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	444	444	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	668	668	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	696	696	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	724	724	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	728	728	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	776	776	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	857	857	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	880	880	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	890	890	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	899	899	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	927	927	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1056	1056	N-LINKED (GLCNAC. .) (POTENTIAL).
SQ	SEQUENCE	1163 AA;	128343 MW;	A7A30784898232F CRC64;

Query Match 25.8%; Score 1516.5; DB 1; Length 1163;
Best Local Similarity 33.9%; Pred. No. 2.6e-94;
Matches 396; Conservative 216; Mismatches 466; Indels 89; Gaps 36;

QY	1	ENLDTENAMTFQENA-RGFGQSVVOLQGSRVVVGAPQBIVAANQORSLVQCDYSTGSCRP	59
DB	24	YMLDTRPTQSFIAQGRHFGYQVLQEDG-VVVGAPGE---GDTGGLTACTSTSSFCQP	79
QY	60	IRLQVEVAVNMSLGLSLAATTPPOLLACGPTVHQTCSNTYVVKGLFCPLFSNLRQQPQ	119
DB	80	VSLH-CSNHTSKYLGMTLATDAAGSLACDGLSRTCDQNTYLSGLYLPQSLGPMML	138
QY	120	KPEALRGCFQEDSDIARLIDGSGTIIPHDFRMKELVSTIMEQLKSKTSLFSLMOYSEE	179
DB	139	QMRPAYQECMKGVLDVLFDPGDSQSLDRKDFKILFEMCDVMVKLSNTSYQFAAVGSTD	198
QY	180	FRIHTFKEF-QNNPNRSLIKPIITQLLGRTHATGLRKVVRLEFNITNGARKNAFKILF	238
DB	199	CRTEPTFLDYVQKNKFDVLLGSVQPMPLLTNTFRATVYVAHVFKESGARPDAFKVLV	258
QY	239	LLTDGEKP--GDPLGYEDVPELDREGVIRYVLGRGDAFRSEKSRQELNTVASKPPRDHV	296
DB	259	IITDGEASDKNISAAHD-----ITRYIIGIKHFVSVQKOKTLHFASEPVEEFV	309
QY	297	FQANNPEALKTVQNLREKIFAIEGTQTGSSSSFFHEMSQEGFSAAITNGPLSTVGSY	356
DB	310	KILDATPEKLKDLFTDLQRRIVAIEGTNRQDLTSPNNELSSSGISADLSKGHAVVGAVGAK	369
QY	357	DWAGGVF-LYTSKEKSTFINMTRVDSMDNDAYLGAA-ILLRNVQSLVLCAPRYOHLG	414
DB	370	DWAGGFLLDREDLQAGATFVQGEPLTSDVRGGYGLTVVAMTSSRSRPRLLAAGAPRYQVG	429
QY	415	LVAMPR--QNTGMBWBSNANVKGTOIGAYFGASLCSDVDVDSNGSTDLVIGAPHYEQTRG	472
DB	430	QVLLFQAPGAGRWQTKLEGTQIGSYFGGELCSVDLDQDGEAEILLIGAPLFPGEQRG	489
QY	473	QGVSVCPILPRQGRARWQCDVLYGQGGQPMWRFGAALTVLGVNGDKLTDVAIGAPGED	532
DB	490	GRVFTY--QRRQSLPEMVSEIQGGPGYPLGRFGAALTALTDINGDLRTDVAVGAPLEP-	545
QY	533	NRGAVYLPHTGSGSGISPSHSORLAGSKLSRLQYFGOSLSGGQDLTMDGLVDLTVAQOG	592
DB	546	-QGAVYIFNGKFG-GLSPQPSORIQAQVFPGRIRFGRSIHVVKDLGGDLADVVVGAG	603
QY	593	HVLLRSQPLRVKKAIMEFNPREVARNVFECNDQVYVKGKAG--SVRVCLHVQKSTRDLR	651

Db 604 RVVLSRPVVDVVTLSFSPFIPVHEVECSYAREQKHGKVKAKFIKILITQ--F 661
Qy 652 EGQIQSVVTVLDALDSORPHRAVENETKMTSTRQTOVLGTCTETKLOLQFNCEIDPV 711
Db 662 QGRLANLSTYLDQGHMSRGLFPDGSHELSTGTSITP-DKSCLDLPHFPHICIQDLI 720
Qy 712 SPIVLRNFSIV---GTPLSAFEN-LRPVLAEDAQRLLFTALFPFKNGMDNLCODLSI 767
Db 721 SPINVSINFSLEBEGTIPROKGRAMPILRPSHTV-TKEIFPKNGCBKDCENLIL 779
Qy 768 TFSFMSLDCLVWGP-----REFNVTVTVVRNDEDSYRTQVTFPPFLDLISYRKVSTLQN 821
Db 780 SSPARS-----GPLRLSSASLAVENWTLNSGEDAYVVRLLDLPRLGSPKRVEMLQ- 831
Qy 822 QRQRORWRLACESASSSTEVSAL-KSTSCSINPIPEPENSEVTNITFDVDSKASLGK 880
Db 832 --PHSRMPVSCBELL--TEGSLTKLKCNVSPPIFKAGQEVSLQVWNTLILNSWEDFV 887
Qy 881 LKXNVISEN-NMPRTNKTETFOLELPVKYAVYVTSHTGVSTKYLAFNTASENTRVWQH 939
Db 888 ELNGTVHCENENSLQEDNSAATHIPVLYPWNILTKQENSTLYISFTPKGPTQVQHV 947
Qy 940 QVSNLQORSILPISLFLVPLVRLNQTWINDRPO-----VTFSENLSS-----TCHTKZ-R 990
Db 948 YQV-----RIQPSAYDHNMT-LBALVGVPRPHSEDLITYTWSVQTDPLVTCHESE 1001
Qy 991 SHSDPLAELKAPVYVNCIAVQCRIQDIFFPGIQEENATLKNLSFDWYIKTSHNLL 1050
Db 1002 SSB---AEQPCFQV-----QFRCPIV---RWEILLQVGTVELSKELKAS-STLS 1046
Qy 1051 IVSTAELFNDSTVTLPGQAFVRSQETKVPFVFNPLPIVSSVGLLLALLITA 1110
Db 1047 LCSLSVSFNSSRPHLYGSKA-SEAQVLVKVBLIHEKEMLVVVLGGLVLLLEFL 1105
Qy 1111 ALYKLGFFKQYKQDM-SEGPPGAEP 1136
Db 1106 ALYKVGFFKELKEKEADGGVNGSP 1132

RESULT 7

ITAE MOUSE STANDARD; PRT; 1167 AA.
AC Q60677;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-E precursor (Integrin alpha M290).
GN ITCAS.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AKR;
RX MEDLINE=9518792; PubMed=7882170;
RA Kishaw P.J., Weis J.H.;
RA "Murine M290 integrin expression modulated by mast cell activation.";
RT Immunity 1:393-403(1994).
CC -1- FUNCTION: INTEGRIN ALPHA-E/BETA-7 IS A RECEPTOR FOR E-CADHERIN. IT
CC MEDIATES ADHESION OF INTRA-EPITHELIAL T-LYMPHOCYTES TO EPITHELIAL
CC CELL MONOLAYERS. MICE EXPRESSING A NULL MUTATION OF THE ALPHA-E
CC SUBUNIT GENE EXHIBIT A MARKED REDUCTION IN THE NUMBERS OF
CC INTRAEPITHELIAL LYMPHOCYTES IN THE GUT AND IN THE DEVELOPMENT OF
CC GUT-ASSOCIATED LYMPHOID AGGREGATES, SUPPORTING A SPECIFIC ROLE FOR
CC THIS INTEGRIN IN MEDIATING RETENTION OF LYMPHOCYTES IN THE
CC INTESTINAL WALL.
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA
CC SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAINS LINKED BY A
CC DISULFIDE BOND. ALPHA-E ASSOCIATES WITH BETA-7.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A WFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -1- SIMILARITY: Belongs to the integrin alpha chain family.
CC -1- SIMILARITY: Contains 1 WFA domain.
CC -1- SIMILARITY: Contains 7 FG-GAP repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U12236; AAC52142.1; --
CC HSSP: P11215; LA8X.
CC MGD; MGI:1298377; Itgae.
CC InterPro: IPR000413; Integrin_alpha.
CC InterPro: IPR002035; VWF A.
CC Pfam: PF01839; FG-GAP; 3.
CC Pfam: PF00357; Integrin_A; 1.
CC Pfam: PF00092; vwa; 1.
CC PRINTS: PRO1185; INTEGRINA.
CC PRINTS: PRO0453; VWFADOMAIN.
CC SMART: SM00191; Int_alpha; 3.
CC SMART: SM00327; VWA; 1.
CC PROSITE: PS00242; INTEGRIN_ALPHA; 1.
CC PROSITE: PS00234; WFA; 1.
CC Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
CC Signal; Repeat; Magnesium;
CC Calcium.
CC SIGNAL 1 19
CC CHAIN 20 1167 BY SIMILARITY.
CC CHAIN 20 181 INTEGRIN ALPHA-E.
CC CHAIN 183 1167 INTEGRIN ALPHA-E LIGHT CHAIN.
CC DOMAIN 20 1114 INTEGRIN ALPHA-E HEAVY CHAIN.
CC TRANSMEM 1115 1137 EXTRACELLULAR (POTENTIAL).
CC DOMAIN 1138 1167 POTENTIAL.
CC REPEAT ? ? CYTOPLASMIC (POTENTIAL).
CC REPEAT ? ? FG-GAP 1.
CC REPEAT ? ? FG-GAP 2.
CC REPEAT ? ? X-DOMAIN (EXTRA DOMAIN).
CC REPEAT ? ? VWA.
CC REPEAT ? ? FG-GAP 3.
CC REPEAT ? ? FG-GAP 4.
CC REPEAT ? ? FG-GAP 5.
CC REPEAT ? ? FG-GAP 6.
CC REPEAT ? ? FG-GAP 7.
CC CA_BIND 514 522 POTENTIAL.
CC CA_BIND 578 586 POTENTIAL.
CC CA_BIND 646 654 POTENTIAL.
CC DOMAIN 185 191 GLU-RICH (ACIDIC).
CC SITE 1140 1144 GFFKR MOTIF.
CC DISULFID 72 83 BY SIMILARITY.
CC DISULFID 130 164 BY SIMILARITY.
CC DISULFID 698 754 BY SIMILARITY.
CC DISULFID 814 820 BY SIMILARITY.
CC DISULFID 884 898 BY SIMILARITY.
CC DISULFID 998 1023 BY SIMILARITY.
CC DISULFID 1031 1047 BY SIMILARITY.
CC CARBOHYD 51 51 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 256 256 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 314 314 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 341 341 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 364 364 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 418 418 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 437 437 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 718 718 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 773 773 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 829 829 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 846 846 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 911 911 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 925 925 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 968 968 N-LINKED (GLCNAC. .) (POTENTIAL).


```
FT CARBOHYD 1013 1013 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1055 1055 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1066 1066 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1167 AA; 128993 MW; 83331C115DCCCFD CRC64;

Query Match
Best Local Similarity 28.68; Pred. No. 1.3e-68;
Matches 353; Conservative 213; Mismatches 462; Indels 205; Gaps 43;

Qy 1 FNLDTEA--MTPOENARGFQSVQVQGGRRVVVGAPOEIVANQVRS-----LYQCOY 52
Db 20 FNMVDNMAVATALQFGAPAVLSLLHLDPSN-----NQTCLLVARRSSNNTAALYRCAL 74
Qy 53 STGCEBIRLQVPEAVNMSLGLSLAATT--SPQLLAC-GPTVHOCTSENTYKGLCEL 109
Db 75 SI--SPDIACO-PVEHICMPKRGYQVTLVGNHNGVLVLCVQVQARKPSLNSLGTACSL 132
Qy 110 FGSMLRQOPKPPALRG-----C-----PQE 131
Db 133 LTPNLDLQAAYFSDLEGFLDPGAHVDSGDYCRSGKGTGEEKSARRRRTVBEDEED 192
Qy 132 DSDIATFLDGSGLIIPHDPRMELVSTIMEQL--KSKTFLSLMOYSEPRHPTKEF 189
Db 193 GTEIATVDDSGSLGSPDFQAKAFISTMMNPFYKFCFNCNFAVQVGAIVQTEFLQES 252
Qy 190 QNNPNRSLIKPIITQLGRTHRTATGLKRVVRELFNITNGARKNAFKILFLITDGEKFGDP 249
Db 253 RDINASLAKVQSIQVQKVTIKTAGAMQVLDNIFIPSGRSGKALKVMVVLTDGDFGDP 312
Qy 250 LGYEDVTPELDREGVIVYVFGDAPRSEKSRQELNTVASKPRDHPVQANNFEALKTQV 309
Db 313 LNTVTINSQKQVGVFAJGVGRPKNNYRELKLIASDPKBAHTFKVTNTYSALDGLL 372
Qy 310 NQREKIFAIEGTGTGSSSFEHMSQEGFSAITNSGP--LLSTVQSGYDAGGVFLY-TS 367
Db 373 SKLQRIVEHGT---VGDALQYLAQTGFSAQLDKGVLLGTVGAFNWSGGALLYSTQ 429
Qy 368 KEKSTPINVT-RVDS-DMNDAYLGYAAAIILNRVQSLVIGAPRYQIHGVAMFRONTGM 425
Db 430 NGRCCFLMOTAKESRVTQVSYLGYSLAVLRKAGHSYVAGAPRRKLRGAVFLRKEDE- 488
Qy 426 WESNA--NNKGTQIGAYFGASLCSDVDNSNGSTDVLIGAPHYEYTCRGQVSVCPPLR 482
Db 489 -EDAPVRRIEGHWGSGYFSGVLCVDIDMDGTTDFLLVAAPFHINGEGRVYVYQVPE 547
Qy 483 GQARWOCDAVLVGEQCPWGRFGAALTVLGDVNGDKLTQVAGAP-----GEENRGA 536
Db 548 -QDASFSLAHTLSGHPGLTNSRFQFAMAAGVDINQDKFTDVAICAPLEGFGAGDGSYGS 606
Qy 537 VYLFHGTSGSGISPSHSORZAGSKLSPLOYFGSLGGODLTMDGLVLTGCAQGHVLL 596
Db 607 VYINGHSG-GLYDSPQQIETASVASGLHYFGMSVGGDLDFNGDGLADITVSRSAVV 665
Qy 597 LRQPVLRVKAIMEFNPVARNVFECDQVVKGEAGEVRVCLHVQKS---TRDLREG 653
Db 666 LRSPVVDLTVSMFTF-----DALPMVFIGK--DVNLCFEVDSVVVASEPGLREM 715
Qy 654 QIQSVWYTDALDGRHSRAVFNETNSRQTVGLTQTC-----696
Db 716 FLNFTVDVIV-----TKQRQLQCESSGSCCLRKWNGSGFICEHFWLI 760
Qy 697 ETLLQLPNCIBDPVSPVILNLFSLVGTSLAFGNLR-----PVLAEDAQLRFLALF--P 750
Db 761 STEEL-----CEEDCFSNITIKVTE-----PQTSGGRRDYPNPTL--DHYKEPSALFQLP 809
Qy 751 FEKNCNDNTIQDGLSITFSFMSLDCLVVGPRFNVTVTVRDGDSYQTQVTFPFLD 810
Db 810 YEKDCKNKVFCIAIEIQTU--ISQELVGVTVKEVTWNISLITNSGDSYNTNMAIYPRN 868
Qy 811 LSVKSVTLQVQSRQSRWLACASASSTEVSGALKTSCTSIINEPIPFENSEVFNTFDV 870
Db 869 IQFKKI-----QKPSVQCCDDPKV-----ASVLWNCKIGHPL-KRSSVNVSVTQWL 918
```

RESULT 8

ITAE_HUMAN STANDARD; PRT; 1179 AA.

AC P38570; Q9NZU9;

DT 01-OCT-1994 (Rel. 30, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Integrin alpha-E precursor (Mucosal lymphocyte-1 antigen) (HML-1 antigen) (CD103 antigen) (Integrin alpha-IEL).

GN ITGAR.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]_

RP SEQUENCE FROM N.A., AND SEQUENCE OF 19-38 AND 179-188.

RC TISSUE=Leukemia, and Lymphocytes;

RX MEDLINE=94164962; PubMed=811947;

RA Shaw S.K., Cepek K.L., Murphy E.A., Russell G.J., Brenner M.B., Parker C.M.;

RT "Molecular cloning of the human mucosal lymphocyte integrin alpha E subunit. Unusual structure and restricted RNA distribution.";

RL J. Biol. Chem. 269:6016-6025(1994).

RN [2]

RP REVISIONS TO 88-114.

RA Parker C.M.;

RA Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE OF 53-1179 FROM N.A.

RC TISSUE=Fetal kidney;

RX MEDLINE=20138496; PubMed=10673275;

RA Touchman J.W., Anikster Y., Dietrich N.L., Maduro V.V., McDowell G., Shotelersuk V., Bouffard G.G., Beckstrom-Sternberg S.M., Gahl W.A., Green E.D.;

RT "The genomic region encompassing the nephropathic cystinosis gene (CTNS): complete sequencing of a 200-kb segment and discovery of a novel gene within the common cystinosis-causing deletion.";

RL Genome Res. 10:165-173(2000).

RN [4]

RP MUTAGENESIS OF ASP-109 AND PHE-316.

RX MEDLINE=20400502; PubMed=10837471;

RA Higgins J.M.H., Cernadas M., Tan K., Irie A., Wang J.-H., Takada Y., Brenner M.B.;

RT "The role of alpha and beta chains in ligand recognition by beta 7 integrins.";

RL J. Biol. Chem. 275:25652-25664(2000).

CC -!- FUNCTION: INTEGRIN ALPHA-E/BETA-7 IS A RECEPTOR FOR E-CADHERIN. IT MEDIATES ADHESION OF INTRA-EPITHELIAL T-LYMPHOCYTES TO EPITHELIAL CELL MONOLAYERS.

CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAINS LINKED BY A DISULFIDE BOND. ALPHA-E ASSOCIATES WITH BETA-7.

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: EXPRESSED ON A SUBCLASS OF T-LYMPHOCYTES KNOWN
CC AS INTRA-EPITHELIAL LYMPHOCYTES WHICH ARE LOCATED IN MUCOSAL
CC EPITHELIAL CELLS.
CC -1- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -1- SIMILARITY: Belongs to the integrin alpha chain family.
CC -1- SIMILARITY: Contains 1 VWFA domain.
CC -1- SIMILARITY: Contains 7 FG-GAP repeats.
CC -1- DATABASE: NAME=PROW; NOTE=CD guide CD103 entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cdi03.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L25851; AAB59359.2; -
CC EMBL; AF168787; AAP43107.1; -
CC PIR; A53213; A53213.
CC HSP; P11215; IABX.
CC Genew; HGNC:6147; ITGAE.
CC MIM; 604682; -
CC GO; GO:0008305; C:integrin complex; TAS.
CC GO; GO:0004895; F:cell adhesion receptor activity; TAS.
CC InterPro; IPR000413; Integrin_alpha.
CC Pfam; PF01839; FG-GAP; 3.
CC Pfam; PF00357; integrin_A; 1.
CC Pfam; PF00092; vwa; 1.
CC PRINTS; PR01185; INTEGRINA.
CC PRINTS; PR00453; VWFA DOMAIN.
CC SMART; SM00191; Int_alpha; 3.
CC SMART; SM0327; VWA; 1.
CC PROSITE; PS00242; INTEGRIN_ALPHA; 1.
CC PROSITE; PS0234; VWFA; 1.
CC Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
CC Signal; Repeat; Polymorphism; Magnesium; Calcium.
CC SIGNAL 1 18
CC CHAIN 19 1179 INTEGRIN ALPHA-E.
CC CHAIN 19 1177 INTEGRIN ALPHA-E LIGHT CHAIN.
CC CHAIN 179 1177 INTEGRIN ALPHA-E HEAVY CHAIN.
CC DOMAIN 19 1124 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 1125 1147 POTENTIAL.
CC DOMAIN 1148 1179 CYTOPLASMIC (POTENTIAL).
CC DOMAIN 181 198 GLU-RICH (ACIDIC).
CC REPEAT ? ? FG-GAP 1.
CC REPEAT ? ? FG-GAP 2.
CC DOMAIN 145 199 X-DOMAIN (EXTRA DOMAIN).
CC DOMAIN 200 391 VWFA.
CC REPEAT 401 456 FG-GAP 3.
CC REPEAT 457 506 FG-GAP 4.
CC REPEAT 510 571 FG-GAP 5.
CC REPEAT 573 638 FG-GAP 6.
CC REPEAT 641 693 FG-GAP 7.
CC CA BIND 522 530 POTENTIAL.
CC CA BIND 586 594 POTENTIAL.
CC CA BIND 654 662 POTENTIAL.
CC SITE 1150 1154 GREY MOTIF.
CC DISULFID 70 79 BY SIMILARITY.
CC DISULFID 126 159 BY SIMILARITY.
CC DISULFID 706 762 BY SIMILARITY.
CC DISULFID 823 829 BY SIMILARITY.
CC DISULFID 893 907 BY SIMILARITY.
CC DISULFID 1008 1033 BY SIMILARITY.
CC DISULFID 1041 1057 BY SIMILARITY.
CC CARBOHYD 49 49 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 321 321 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 444 444 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 726 726 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 782 782 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 857 857 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 934 934 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 954 954 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1065 1065 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1096 1096 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 360 360 D > E.
FT VARIANT 1041 1041 /FTID=VAR_008884.
FT VARIANT 208 208 /FTID=VAR_008885.
FT MUTAGEN 316 316 D->A: LOSS OF B-CADHERIN BINDING.
FT MUTAGEN 477 477 F->A: LOSS OF B-CADHERIN BINDING.
FT CONFLICT 482 482 V->I (IN REF. 3).
FT CONFLICT 482 482 Q->R (IN REF. 3).
FT CONFLICT 950 950 R->W (IN REF. 3).
FT CONFLICT 1019 1019 A->V (IN REF. 3).
SQ SEQUENCE 1179 AA; 130088 MW; E558902EDF9D95E1 CRC64;
Query Match 19.2%; Score 1128; DB 1; Length 1179;
Best Local Similarity 28.7%; Pred. No. 5,1e-68;
Matches 338; Conservative 213; Mismatches 454; Indels 172; Gaps 39;
QY 45 GSXYQDYSTGS--CEPI-RLQVP-----VEAVNMSGLSLAATSPOLLACGPTVHQ 95
DB 65 GPLHRCSLVQDEILCHPVEHPVPRGRHGVTVVRSHHGVLICI-----QVLVRP--HS 117
QY 96 TCSENYVYKGLCFLPGSNLRQPO-----SDAFIDGSGSIIPHDPRMKELVSTIMEQL-- 119
DB 118 LSELT---GTCSLLGPDRLRPOAQNPFLENLDFDARVDTGDCYSNKEGEGDDVNTA 174
QY 120 KFEALRGCPQED-----SDFIADGSGSIIPHDPRMKELVSTIMEQL-- 164
DB 175 RQRRALEKEEEDKEEEDBEAEAGTEIATLDGSGSIDPDPDFORAKDFISNMVNFYE 234
QY 165 KSKSTLESIMQYSEEFRIHFTFKFQNNPNRSLKPTOLLGRTHTATGLKVVRELEN 224
DB 235 KCFECNFALVQYGVQTEFLDLRSDQWASLARVNITQVGSVTKTASAMQVLDOSIF 294
QY 225 ITNGARKNAFKILFLITDGEKFGDPLGYEDVPELDREGVIRYVLGFGDAPSEKROEL 284
DB 295 SSHGSRKASKVMVLTDDGIFEDFLNLTTVINSPKMQGVVERFAIGVGEFSAKTAREL 354
QY 285 NTVASKPPDRHVQANFEALKTQVQOLREKIPALECTCTGSSSSSPEHEMSQEGSAAT 344
DB 355 NLIASDPDDETHAFKVTNTYALDGLSLKLYNIISMEGT---VGDALHYQLAQIGFSAQL 411
QY 345 SNCP--LLSTVGSYDWAGGVFLY--TSKEKSTFINMTRVDSMDNDA---YLGYAAAILRN 398
DB 412 DERQVLLGAVGAFDWSGGALLYDTSRRGRFLNQTAADAADEAAQYSLGVAVAVLHKT 471
QY 399 RVQSLVLCAPRYOHI GLVAMFR--QNTGWESUNVY--KGTQICAYFGASLCSVDVDSNGST 456
DB 472 CSLSYVAGAPQYKHG--AVFELQKEGREASPLVLEGEQMGSYFGSELCPYDIDMDGST 529
QY 457 DLVLIGAPHYTEQTRGGQVSVCPPLRGORARQCDVAVLYGEOQPMWRFGAALTJVLGVN 516
DB 530 DFLVAAPFYVHVGEEGRVYVRLSE--QDGSFSLARILSGHPGFTNAPFGFMAAGDLS 588
QY 517 GDLTADVAIGAP---GREDNR--GAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFQ 570
DB 589 QDKLTADVAIGAPLEGFGADDOGASFGSVIYNG--HWDGLSASPSQIRASTAVAPGLQYFG 647
QY 571 SLSCGQDLTMDGLVDLTVGAQGHVLLRSPQLRVKATMEFNPREFARNVFECDQVVK 630
DB 648 SMAGFDISGGLADITVGTLGQAVFRSRPVVRUKVSKAFTPSALP-----ICF 697
QY 631 KEAGRVVRLCHVQKSTRDLREGQIQSVVTVYDALDLSGRPHSRFVFNFTKSTRQTVL 690
DB 698 NGVNVRLCFEI--SSVTTASBSGLREALNPLTDLVDVGKRRRLQCCSDVRSCLGLREWS 756
QY 691 GLUTQCEVILKQLPN-----CIEDPVSPIVRLNFSVGTFLSAFGLNPLVLAEDARLPT 746

```
Db 757 SCSQCELDL-LMPTGELCEBDCPNSASVKVSYQL-QTPGQTDHPQILDRIYTEPPAI 814
Qy 747 ALPFFKNCNDNIQDDLSITFSFMSLDCLVVGGPRENVTVTVRNDGEDSVRTQVTFP 806
Db 815 FQLPYEKACNKLFCVAELQLA-TTVSQCELVVGLTKELTLNLTNSGEDSVYTMALN 873
Qy 807 FPLDLSVRKYSTIQMORSQSRWELACESASSTEVSGALKSTSCSINHPIPPENSEVTENI 866
Db 874 YPRNLQ-----LKRQKPSFNIQDDPPQV---ASVLIMNCRIGHPVL-KRSSAHVSV 923
Qy 867 TPVDSDKASLGNKLLKANVTSENN---MPRTNKTEFQ---LELPVKVAVYVTVTSHG 919
Db 924 VMQLEENAPNRTADITVTVTNGNRRSLANETHTLQFRHGFVAVLSKPSIMYVNTGQGL 983
Qy 920 S-TKVLNFTASENTSEVMQHOVQVSNLQORSIPISIVPLNVRINQTVIMDRPQVTFSE 977
Db 984 SHKGFELFHVHGEN---LFGAEYQ-----LQICVPTKRLGLQVAVKKLRITQ 1028
Qy 978 NLASTCHTKERLPSHSDFLAELRKAPVYVNCISAVCQRIQCDIPFFGIGQEBFNATLKNLS 1037
Db 1029 ASTVCTWSQERACAYSS-VQVHEEHSVSCVIA-----SDKENVTVAABIS 1073
Qy 1038 PDWYIKTSHNHLIVST-----ABTLFNDSTVTLPGQCAFVRSQTSKVEPFPVNP 1091
Db 1074 WD-----HSEELKDVTELQILGEISFNKSLYEGNAENH--RTKITVTVFLKDKYHSL 1125
Qy 1092 PLVIGSVGGLLLALITAAALYKLGPFKROYKDMNSE 1128
Db 1126 PIILKSGVGLLVILVILVLPKCGFKRYKQQLNLE 1162

RESULT 9
ID ITAL HUMAN STANDARD; PRT; 1151 AA.
AC P56139;
AT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Integrin alpha-1 (laminin and collagen receptor) (VLA-1) (CD49a).
GN ITGA1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93155124; PubMed=8428973;
RA Briesewitz R., Epstein M.R., Marcantonio E.E.;
RT "Expression of native and truncated forms of the human integrin alpha
FT 1 subunit";
RL J. Biol. Chem. 268:2989-2996(1993).
CC -!- FUNCTION: INTEGRIN ALPHA-1/BETA-1 IS A RECEPTOR FOR LAMININ AND
CC COLLAGEN. IT RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-P-P-G-
CC E-R IN COLLAGEN.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-1
CC ASSOCIATES WITH BETA-1.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- DOMAIN: THE INTEGRIN 1-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH 1-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -!- DATABASE: NAMB=PROW; NOTE=CD guide CD49a entry;
CC WWW=http://www.ncbi.nlm.nih.gov/prow/cd/cd49a.htm.
DR PIR; A45226; A45226.
DR PDB; 1QCS; 17-MAY-00.
DR GeneW; HGNC:6134; ITGA1.
DR MIM; 192968; -.
DR GO; GO:0008305; C:integrin complex; TAS.
DR GO; GO:0004895; P:cell adhesion receptor activity; NAS.
DR GO; GO:0005518; P:collagen binding; TAS.
DR GO; GO:0007160; P:cell-matrix adhesion; NAS.
DR InterPro; IPR000413; Integrin_alpha.
```

```
DR InterPro; IPR002035; VWFA_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWFA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWFA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Repeat; Calcium; Magnesium; 3D-structure.
FT DOMAIN 1 1113 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1114 1136 POTENTIAL.
FT DOMAIN 1137 1151 CYTOPLASMIC (POTENTIAL).
FT REPEAT 16 75 FG-GAP 1.
FT REPEAT 2 2 FG-GAP 2.
FT REPEAT 147 360 VWFA.
FT REPEAT 349 404 FG-GAP 3.
FT REPEAT 405 457 FG-GAP 4.
FT REPEAT 459 520 FG-GAP 5.
FT REPEAT 540 599 FG-GAP 6.
FT REPEAT 602 654 FG-GAP 7.
FT CA_BIND 470 478 POTENTIAL.
FT CA_BIND 552 560 POTENTIAL.
FT CA_BIND 614 622 POTENTIAL.
FT SITE 1139 1142 GFFKR MOTIF.
FT DISULFID 54 64 BY SIMILARITY.
FT DISULFID 660 669 BY SIMILARITY.
FT DISULFID 675 728 BY SIMILARITY.
FT DISULFID 780 786 BY SIMILARITY.
FT DISULFID 850 858 BY SIMILARITY.
FT DISULFID 1002 1034 BY SIMILARITY.
FT DISULFID 1037 1044 BY SIMILARITY.
FT CARBOHYD 46 46 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 72 72 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 77 77 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 84 84 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 189 189 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 289 289 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 313 313 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 374 374 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 432 432 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 504 504 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 671 671 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 720 720 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 752 752 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 812 812 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 855 855 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 880 880 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 887 887 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 911 911 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 938 938 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 946 946 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 980 980 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1045 1045 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1055 1055 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1074 1074 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1085 1085 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1151 AA; 127837 MW; 683F3C1AABF52808 CRC64;
```

```
Query Match 18.5%; Score 1085.5; DB 1; Length 1151;
Best Local Similarity 27.6%; Pred. No. 3.7e-65;
Matches 342; Conservative 209; Mismatches 488; Indels 201; Gaps 44;
Qy 1 FNLDTENAMTFCQENARG-FGQSVVOL---QGSRVVVGAPQETVAANQRGSLYQCDYSTGS 56
Db 1 FNVVDKNSMTFGSPVDEMGYTVQVYEBEGKNWVLIGSLVQCPKRTGDVVKCPVGRGE 60
Qy 57 CEP-IRLQVPEA-----VNSLGLSLAATSPPOLLAGCTVHTGCTSENTYVKGL 106
Db 61 SLPCVKLLDLPVNTSIPNVTVENKNTMFGSTL-VTNPNGGFLACGPLYAYRCGLHYTTGI 119
Qy 107 CFLFGSNLEQQQKPFALRGCFQEDSDIAFLIDSGSGIIPEDFRMKELVSTIMEQLKK 166
```

Db 120 CSDVSTPQVNSIAB--VQECSTQ-LDIVIVLDGNSIYPWD-----SVTAFLNDLLKR 171
Qy 167 -----SKTFLSMQVSEFRIRIHTFKFQFNPNPNSRLIKPITQOLLGR-THETATGLRKVR 220
Db 172 MDIGPKQFQVGVQVGENYTHFNLNKYSSTSEVLVAAKKIVQRGQRTMTALGTDITARK 231
Qy 221 ELFNITNGARKNAKFLILTDGKFGDPLGYEDVPELDREGVIRYVLGFGDAFR---- 276
Db 232 EAFTEARGARRGVKKVMVITDGESH-DNHLRLKVIQDCEDENIQFSTAILGSYNRGNL 290
Qy 277 -SEKRSQELNIVASPKPRDHVFOANNFEALKTQVQNLREKIPAIKGTQGTGSSSEHEHMS 335
Db 291 STEKFVEEIKSIASEPTEHFFNVSDALAVITVKTGLGRIFALEATADQSAASEMEMS 350
Qy 336 QSGFSAIITNGSPLSTGVSVDWAGVFLYTSKE-----KSTF-INMTRVDSMDMDAVLG 389
Db 351 QTFGSNHYSDWMLGAVGAYDNGVTVMKQASQIIIPNTTFNVSTKKEPL-ASVLG 409
Qy 390 YAAAILNRVQSL-VLGPAPYOHIGLVAMFQNTGMESNANVKGTQIGAYFGASLCSV 448
Db 410 YTVNSATASSGDVLYIAGQPRYNTQVYIYRMEDGNIKILOTLGSEQIGSYFGSILITF 469
Qy 449 DVDNNGSTDLVILGAPHY-----YECYR-GGOVSVCPLEPRGQARWQDA 492
Db 470 DIDKDSNTDILLVGPAPMYMTKEBQGVYVVALMQTRFEYQMSLEPIKQTCSSRQHNS 529
Qy 493 VLYGEGQCPWG-RFGAALTGLVDVNGDKLTDVAIGAPGEEDNRGAVYLFPHGTSGSGI 551
Db 530 CTTENKNEPCGARFGTJAAVADKMLDNDGVIVIGAPLEDHGGAVYVHG-SGKTIKRE 588
Qy 552 HQRIAGSKLSPRLQVFGSLSGGQDLTDGLVDLTGVAQGVHLLRSQPLRVKAIMEF 611
Db 589 YAQRPISGGDKTLKPFQSGSIHGMELDNGDGTVDVTIGLGAALFWSRDAVAVKVTWVF 648
Qy 612 NPREVARNVECDQVVKKEAG--EVRVCLHVO-KSTRDLRREGIQSVVYDLDALDSG 668
Db 649 ENKVMIOKNCH--MECKEITVCINATVCFVKLSKEDTIVEADLQ----YRVLDSL 701
Qy 669 RHRSRAVNET-----KNSRTRQVGLGTQTCETLKLQNLNCTEDPVPVILNLSL 722
Db 702 RQISRSFFSGTQERKQVQRTNITVKSEC-----TKHSFYMLDXHDFQDSVR---ITLDFNL 753
Qy 723 VGTPLSAFGLNAPVLAEDAQLFTALFPFEKNGKNDNICODLSLTFSPMSLDCLVVGPP 782
Db 754 T-DPENG-----PVLDDSLPNSVHEIYIPAKCGCKEKCISLSLHVATEKDLIVRSQ 807
Qy 783 RE-FNVTVTRNDGSDSYRTQVFFPPLDLSYKSTLQNRQSRQSWRLACSASTEVS 841
Db 808 NDKFNVLTVKNTKDSAYNTRTIVHYSPLVSPGTEAIQKD-----SCBSN----- 853
Qy 842 GALKSTCSINHIPIFENSEVTFNITFDVDSKASLGN-KLLAKANVTSENMPRTNKTEF 900
Db 854 ---HNITCKVGVFFLRGRBNVTFKILFQNTSYLMENVTIYLSATSDSEPPETUSDNVV 910
Qy 901 QLELPVKYAVVWVTVSHGVSTKYNFTASENTSRVMHQYQVSN-----LGQRS----- 949
Db 911 NTSIPVKYEVGLQFYQ-SASEYHISAETNTEVINSTEDIGNEINIFYLIRKSGSPM 969
Qy 950 ---LPLSLVF-----LVPRLNQTIVDRPQVTFSENLSSTCHYE-----R 988
Db 970 PELKLSISFPNMTSGPYLYPTGLSS-----SENANCRPHIFEDPFSINSKGL 1018
Qy 989 LPSHSDFLAELAKAPVNVCSIAVCQRIQDIPFFGIQE-----EFNA 1030
Db 1019 MTTSTD---HLKRGITLDCNTKCFATITCNLTSSDISQVNVSLILWKTPFKSYFSSLNL 1075
Qy 1031 TLKGNLSDFOWYIKTSRNNHLLIVSTABILFNDSVFTLLPQGGAFVRSQVETKVEPEVNP 1090
Db 1076 TIRGEL-----RSENASLVLSSN-----QKRELAIQISDKGLPGR 1111
Qy 1091 LPL--TVGSSVGLLLALITLALYKLGFFKQYKDMSE 1128

Db 1112 VPLWVILLAPAGLMLLMLLALWKLKIGFFKRPKLKKMEK 1151

RESULT 10

ITAH HUMAN

AC Q9UKX5; Q9UKQ1; STANDARD; PRT; 1189 AA.

DF 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Integrin alpha-11 precursor.

GN ITG11.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

FN [1];

SEQUENCE FROM N.A.

RP TISSUE=Fetal heart, and Osteoblast;

RC MEDLINE=99417678; PubMed=10486209;

RA Lehnert K., Ni J., Leung E., Gough S.M., Weaver A., Yao W.P., Liu D., Wang S.-X., Morris C.M., Krissansen G.W.;

RA "Cloning, sequence analysis, and chromosomal localization of the novel human integrin alpha11 subunit (ITG11).";

RL Genomics 60:179-187(1999).

RN [2];

SEQUENCE FROM N.A.

RP TISSUE=Fetal muscle, and Uterus;

RC MEDLINE=99395147; PubMed=10464311;

RA Velling T., Kusche-Gullberg M., Sejersen T., Gullberg D.;

RT "cDNA Cloning and Chromosomal Localization of Human alpha(11) Integrin. A collagen-binding, i domain-containing, beta(1)-associated integrin alpha-chain present in muscle tissues.";

RL J. Biol. Chem. 274:25735-25742(1999).

RN [3];

SEQUENCE OF 954-1188 FROM N.A.

RP TISSUE=Fibroblast;

RC Andreu N., Estivill X., Escarceller M., Sumoy L.;

RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: INTEGRIN ALPHA-11/BETA-1 IS A RECEPTOR FOR COLLAGEN.

CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-11 ASSOCIATES WITH BETA-1.

CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

CC -!- TISSUE SPECIFICITY: ACCORDING REF.2 HIGHEST LEVELS IN UTERUS AND HEART. INTERMEDIATE LEVELS IN SKELETAL MUSCLE AND INTERMEDIATE TO LOW LEVELS IN PANCREAS, KIDNEY AND PLACENTA. ACCORDING REF.1 ALSO FOUND IN BRAIN, COLON, LUNG, SMALL INTESTINE, STOMACH, TESTIS, SALIVARY GLANDS, THYROID GLANDS AND PROSTATE. VERY LOW LEVELS IN PERIPHERAL BLOOD LYMPHOCYTES, FETAL BRAIN AND FETAL LIVER.

CC -!- DEVELOPMENTAL STAGE: STRONGLY UP-REGULATED IN DIFFERENTIATING FETAL MUSCLE CELLS (IN VITRO).

CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VMFA DOMAIN. INTEGRINS WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.

CC -!- SIMILARITY: Belongs to the integrin alpha chain family.

CC -!- SIMILARITY: Contains 1 VMFA domain.

CC -!- SIMILARITY: Contains 7 FG-GAP repeats.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC EMBL; AF109681; AAF01258.1; -

DR EMBL; AF137378; AAD51919.2; -

DR EMBL; AL359064; CAB94392.1; -

DR HSPG; P17301; IAOX.

DR Genew; 604789; -

DR MIM; 604789; -

DR GO; GO:0008305; C:integrin complex; TAS.

DR GO; GO:0004895; F:cell adhesion receptor activity; TAS.

DR	GO: GO:0005518; F:collagen binding; TAS.	GO: GO:0007160; P:cell-matrix adhesion; TAS.	GO: GO:0007517; P:muscle development; TAS.	InterPro; IPR000413; Integrin_alpha.	InterPro; IPR002035; WVF_A.	Pfam; PF01839; FG-GAP; 3.	Pfam; PF00092; vwa; 1.	PRINTS; PR01185; INTEGRINA.	SMART; SM00191; Int_alpha.5.	SMART; SM00327; WVA; 1.	PROSITE; PS00242; INTEGRIN_ALPHA; FALSE_NEG.	PROSITE; PS0234; WVFA; 1.	Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane; Signal; Repeat; Calcium; Magnesium; Polymorphism.	SIGNAL	1	22	CHAIN	23	1189	DOMAIN	23	1142	TRANSMEM	1143	1165	DOMAIN	1166	1189	REPEAT	38	94	REPEAT	102	163	DOMAIN	167	345	REPEAT	359	420	REPEAT	422	475	REPEAT	477	537	REPEAT	539	598	REPEAT	601	653	DOMAIN	1154	1162	DOMAIN	1174	1177	CA_BIND	488	496	CA_BIND	551	559	CA_BIND	613	621	DISULFID	76	83	DISULFID	121	139	DISULFID	129	159	DISULFID	659	668	DISULFID	674	729	DISULFID	781	787	DISULFID	881	893	CARBOHYD	82	82	CARBOHYD	95	95	CARBOHYD	291	291	CARBOHYD	331	331	CARBOHYD	358	358	CARBOHYD	449	449	CARBOHYD	462	462	CARBOHYD	528	528	CARBOHYD	642	642	CARBOHYD	694	694	CARBOHYD	857	857	CARBOHYD	894	894	CARBOHYD	973	973	CARBOHYD	1032	1032	CARBOHYD	1040	1040	VARIANT	433	433	VARIANT	524	524	VARIANT	972	972	VARIANT	1003	1003	VARIANT	1030	1030	VARIANT	1094	1094	SEQUENCE	1189	AA; 133609 MW; 60303C08A44CD52 CRC64;	Query Match	18.4%; Score 1079.5; DB 1; Length 1189;	Best Local Similarity	28.3%; Pred. No. 9.8e-65;	Matches	349; Conservative 212; Mismatches 501; Indels 173; Gaps 48;
QY	1	FNLDTENAMTFOENARG--FGQSVVO--LQGSR-VVVGAPQEIIVAAQSGSLYCCDYSGS	56	23	FNDTRKPRVIFGSRSTAFFGYTVQOHDISGNKWLWVGAPLETNGYKTDGVDYKCPVIHGN	82	57	CEPIRL-----QVPVEAVNMSLGLSLAATTSPQALLACGPTVHQTCSNNTYVKGICFLF	110	83	CTKMLGRVTLNVSVERKONNRLGLSLATNPKNDSFLACSPWSHSECGSSYVTTGMSRV	142	111	GSNLROQKPPALRGCPQEDSDIAFLIDGSGSIIPHDFRMKELVSTIMEQLK----	166	143	NSNFRPSKTVAP-ALQRC-QTYMDIVI VLDGNSIYP-----WVEVQHPLNLIKPFYTG	195	167	-SKTLFSLMQYSEEFRIHFTFEFQNNPNRSLIKPITQLLG-RTHATGLKRVVRELFN	224	196	PGQIQGVVQYGEDVVEHFLNDYRSVKDVVEAASHIEQRGCTETRTAFGISPARSEAFQ	255	225	ITNGARKNAFKILFLITDGEKFGDPLGYEDVITPELDREGVIRY---VLGPGD--AFRSEK	279	256	--KGRKGAKKVMIVITDGEHSDSP-DLEKVIQOGERDNTVRYAVAVLGYNNRGINPET	312	280	SQELMTVASKPPDRHVFOANNFEALKTVQNLREKIPAEICTQTGSSSSPHEHMSQEGF	339	313	FLNEIKYIASDDDKHFFNVTDAAIKDVIDALGDRIFSLSEGTNK-NETSFGLMSQTFG	371	340	SAATTSNGPLLSTVGSYDWAGGVFLYTSKEK-----STFINMTRVDSMDNDAYLVAAAI	394	372	SSHVEDGVLLGAVGAYDHWGAVLKEYSACKVIPLRESYLKEFPEELKNHGAJLGYTVTS	431	395	IIRNRV-QSLVLAGAPRYOHIGLVAMP-RQNTGMWESNANVKGTQICAYFGASLCSVDVDS	452	432	VYSSRQGRVYVAGAPRFTNHTGKVIILFTMHNNRSLTIHQMRGQOIGSYFGSEITSDIDG	491	453	NGSTDLVLIGAPHYYEQTR-GQGVSVCLPRGQARWQCDVLYGEGQGPWGRFGAALTV	511	492	DGVTDLVVGAPMYFNEGREKGVYVEL---RQNRVYVNGTLKDSHYSQYARFGSSIAS	548	512	LGDVNGDKLTDVAIGAPGEEDNRGAVLPHGTSGSGISPSHSGIAGSKLSPRLVFGQS	571	549	VRDLNQDSYNDVVVVGAPLEDNHAGAIYIFHGFRGS-ILKTPKQRTASLALGLQYFGCS	607	572	LSGGQDLTMDGLVDLTVGAGQHVLLRSQPLRVKAIKMFNPREVARNVF--ECNDQVVK	629	608	IHQQLDLNEDGLIDLAVGALGNVILWSRPVQINASHLFEPSKI--NIFHRDC-----	659	630	GKEAGEVRVCL-----HVQKSTRDLREBQIQSVVYTDIALDSCGPHSRVAF	676	660	-KESGRDATCLAAFLCFTPIFLAPHFTTVG-----IRYNATMDERRYTPRAHL	708	677	NETKNS-TRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRNFSVLVGTPLSAFGLNRP	735	709	DEGGDFTNRAVLSSGQELCERINFHVL-DADYVKKVPTFSVEYSLEDP-----DHGP	761	736	VLABDAQRLFTALPPEKNCNDNICTDDL-----SIT	768	762	MLDDGWPTLRSVFPWNGCNEDEHCVPLVLDARSDLPTAMEYQCVLRKPAQDCSAYT	821	769	PSFMSLDLWVGSPREFNVTVTRNDGDSYRTQVTFPPFLDLSYRKVSTLQNSQSRWS	828	822	LSFDITVPIESTRQVAVAEATLENNGENAYSTVLNISQSANLQF--ASLIQKEDSDGS-	878	829	RLACSASSTVSGALKSTSCSINHPIIPENSEVFTNTTFVDVSKASLGNKLLKANVTS	888	879	-IECVNEER-----RLQKQVCNVSYPFFRAKAKVAPRLDFEF-SKSIHLHLEIELAAGS	931	889	ENNMPTNKTB--FOLELPVKYAVVMVVTSHGVSTKY---LNFTAS--ENTSRVMQHOVQ	941	932	DSNERDSTKEDNVAPLRFHLKYTEADVLFTRSSSLSHYEVKLNSSLYERDYGIFPPFCIFR	991	942	VSNLGO---RSLPISLVFLVPLVRLNQTVIWRDPOVTFSENLSSTC-----HTKERLP	994	992	IQNLGLFPIHGMKTIPIATRSNGNLLKLKDLFT-DEVANTSCNMGNSYRTPVPE	1050	995	FLABLEKAPVNCISIAVCQRIQCDIPFFGIQEEFNATKGNLSFDWY-----IKTSHNL	1049																																																	

Query Match 18.4%; Score 1079.5; DB 1; Length 1189;
Best Local Similarity 28.3%; Pred. No. 9.8e-65;
Matches 349; Conservative 212; Mismatches 501; Indels 173; Gaps 48;

Db 1051 --EDLRARQLNHSNDVSVINCNIPLVP--NOEINPHILGNL---WLSLKALKYKSMKI 1104
QY 1050 LIVSTRAILLFNDVFTLLPQOGAFVRSQTETKVEPEVFN-----PLPLVIGSSVGGILL 1103
Db 1105 MVNAALQROPH-SPF-----IFREEDPSQVFEISKQEDMQVPIIIVGSTLGGILL 1155
QY 1104 LLALITAAALYKLGFFK--ROYKDMSEGGPGGAEP 1136
Db 1156 LLALLVLALWKLGGFFSARRRR-----PGLDLP 1183

RESULT 11
IT2A2 BOVIN STANDARD; PRT; 1170 AA.
AC P53710;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-2 precursor (platelet membrane glycoprotein Ia) (GP1a)
DE (Collagen receptor) (VLA-2 alpha chain) (CD49b) (Fragment).
GN ITGA2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Sutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RX MEDLINE=94193647; PubMed=7511592;
RA Kamata T., Puzon W., Takada Y.;
RT Identification of putative ligand binding sites within I domain of
integrin alpha 2 beta 1 (VLA-2, CD49b/CD29).
RL J. Biol. Chem. 269:9659-9663(1994).
CC -I- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A RECEPTOR FOR LAMININ,
COLLAGEN, COLLAGEN C-PROPEPTIDES, FIBRONECTIN AND E-CADHERIN. IT
RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-P-G-B-R IN
COLLAGEN. IT IS RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER
CELLS TO COLLAGENS, MODULATION OF COLLAGEN AND COLLAGENASE GENE
EXPRESSION, FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED
EXTRACELLULAR MATRIX.

CC -I- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-2
ASSOCIATES WITH BETA-1.
CC -I- SUBCELLULAR LOCATION: Type I membrane protein.
CC -I- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -I- SIMILARITY: Belongs to the integrin alpha chain family.
CC -I- SIMILARITY: Contains 1 VWFA domain.
CC -I- SIMILARITY: Contains 7 FG-GAP repeats.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL; L25886; AAB59255.1; -;
DR HSP; P17301; IAOX.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF A.
DR Pfam; PF01839; FG-GAP; 3-
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR SMART; SM00191; int_alpha; 5.
DR SMART; SM00327; VWFA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWFA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Platelet; Signal; Repeat; Polymorphism; Calcium; Magnesium.
FT NON_TER

FT SIGNAL <1 18
FT CHAIN 19 1170
FT DOMAIN 19 1121
FT TRANSMEM 1122 1143
FT DOMAIN 1144 1170
FT REPEAT 34 92
FT REPEAT ?
FT DOMAIN 177 367
FT REPEAT ?
FT REPEAT 423 475
FT REPEAT 477 538
FT REPEAT 540 599
FT REPEAT 604 656
FT CA_BIND 488 496
FT CA_BIND 552 560
FT CA_BIND 616 624
FT SITE 472 474
FT SITE 1146 1150
FT DISULFID 72 81
FT DISULFID 669 726
FT DISULFID 778 784
FT DISULFID 854 865
FT DISULFID 1008 1039
FT DISULFID 1044 1049
FT CARBOHYD 94 94
FT CARBOHYD 101 101
FT CARBOHYD 332 332
FT CARBOHYD 421 421
FT CARBOHYD 449 449
FT CARBOHYD 464 464
FT CARBOHYD 688 688
FT CARBOHYD 748 748
FT CARBOHYD 945 945
FT CARBOHYD 1063 1063
FT CARBOHYD 1070 1070
FT VARIANT 580 580
FT VARIANT 588 588
FT VARIANT 725 725
SQ SEQUENCE 1170 AA; 128929 MW; 8BCE1C5F2448FB1 CRC64;

Query Match 18.2%; Score 1071; DB 1; Length 1170;
Best Local Similarity 27.5%; Pred. No. 3.6e-64;
Matches 334; Conservative 218; Mismatches 495; Indels 168; Gaps 48;

QY 1 ENLDTENAWTQ-ENARFGQSVVQL---QGSRRVVVGAPOEIVANQSGSLVQC--DYST 54
Db 19 YVGLPKAKIFSGPSEDFGVAVQQFIPNKGWLLVGSFPGKRMGDVYKCPVDIST 78
QY 55 GSCEPIRLQ-----VPEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTVVKGLC 107
Db 79 TTCEKLNLOTSTSMNSNVTMTNMSLGLTLRNVTGGFLTCGLPWAQCGSQYVTTGVC 138
QY 108 FLFGSNLQOQKPEALRGCPQEDSDIAFLDQSGSIIPHDPRMKELVSTIMEQLK-- 165
Db 139 SDVSPDF-QLRTSFAPAVQTCP-SFIDVVVVDENSNISYPWD--AVKNFLEKFGVGLDIG 194
QY 166 KSKTLFSLMQYSSEPRIFHTFKBFQNNPNRSLIKPIITQLL----GRTHATGLKVVRE 221
Db 195 PTKQMGLIQYANNPRVFNLTFKSKD--EMIKATSQTFQYGGDLNTPKAIQVARDT 251
QY 222 LFNITNGARKNAFKILFLTLTDEKFGDPLGVEDVPELDREGVIRY---VLGF--GDAFR 276
Db 252 AYSTAAGGRPGATKVMVVVYTDGESH-DGSKLKAVIDQCNDKONILFAGVILGYLNRNLD 310
QY 277 SEKSQELNTVASKPPRDHVFQANNFEALKTVMQNLREKIFAIEGTQTGSSSSPHEMSQ 336
Db 311 TKQLIKEIKATASIPTEHFPNVSDADLLEKAGTIGQIFSEGTVOG-GNFMQMSQ 369
QY 337 EGFSAAIT--SNGPLLSVGVSDNAGGVFLTSKSKTSPIFIMNT--RVSDNMN-DAYLVGA 391
Db 370 VGFSAEYSPQNNILMLGAVGAYDSGTVQKTPHGHILFYSKQAFQIQLQDRNHSYLYGS 429
QY 392 AAILLRNRVQSLVLCAPRYQHIGLVAMFRQNTGMWESNANV-----KGTQIGAYFGASL 445

```
Db 430 VASISTGNSVHFAAGPRANTGQIVLYSN-----ENGWTVVTSQSGQIGSYFGSVL 484
Qy 446 CSVDVDSKSGTDLVLICAPHYEYQTR--GGQVSVCLPRGORARQWCDVLYGQGGPWG 503
Db 485 CAVDVNKDTITDVLVAGAPMTMDLKKKEGRVYLFTITKG-ILNWH--QFLEGPNGLENA 541
Qy 504 RFGAALTIVLGVNDKLTDAVAGPGRDNGRGAVALFHGTSGSISPSHSQRIAGS--KL 561
Db 542 RFGSAIATLSIDINWGFNDVIVGSPLENQNSGAVIYNGHEGM-IRLRYQKILGSDRAP 600
Qy 562 SPRLOVYFQSSLSGGQDLTMDGLDVLTVGAQGHVLLRSQPVLRKAIEMFNPREVARVP 621
Db 601 SSHLOVYFGRSLDGYGLDNGSDITDVSVCAPGVQLAWSQISADVSVDASFTPKKI--TL 658
Qy 622 ECNDQVVKKZAGEVRVCLVYKSTRDLRGQIQSVVTVYDLAD----SGRPHRAVEN 677
Db 659 KKNABE-----KKLCF-----SAKFRFTQNNQVRAIVNITIDBQFSRSVLSRLFK 707
Qy 678 ETKNSTRTQVLGLTQCE--TLKLQFPNCIEDVPSPVILRLNFSL--VGTPLSAFGLN 733
Db 708 ENNERCLQKTMVQAQRCSFVYIHIQEPS---DIISPLNLCMVISLENPOT----- 756
Qy 734 RPLVLAQRLFTALPPEKXCGNDNICQDLSITP-----SPMSLDCLVVGPREFNTV 789
Db 757 NPALVAXSETVYKVSIFPHKCGDGVGCSIDLVLNVOQLPATQCOQPFIVSNQKRLTFSV 816
Qy 790 TVRNDGDSYRTVTFPPFLDLSYRKVSTLQNRQSRWLRACSSAST-EVSGALKSTS 848
Db 817 QKXKKEAYNTEIVDSENLFF-----ASMSMPVDGTEVTCQIASQKSVT 864
Qy 849 CSNIHPPIPPENSEVTFNTPVDVSKASLGKLLKANVTSMNPRNTKTFQLELPVKY 908
Db 865 CNVGPALKSKQQTTFNFDNLQ-NLQNASISFRALSBSQENMADNSVNLKSLLY 923
Qy 909 AVYVMTVSHGVSTKVLNFTASNTSRVMOHYOVSNLQZ-----SLPISLVFLV 958
Db 924 DAEIHT--ASTNIPYVSLDGNSSV--HSEF--DIGPKPIFKITVTSVPSVMA--- 976
Qy 959 PVRLNQTVIMDRPQVTFSEN---LSSCTHTKE-----RLPSHSDFTLAE- 998
Db 977 -----SVIIHPQVTKDKNPLVLTGVHTDQAGDISCEABINPLKIQOTSSVSFKSEN 1030
Qy 999 LKAPVNCSTAVCORICDIPFGIQEFNATLKNLSFDWYIKTSHNHLIYSTAEI- 1057
Db 1031 FRHIKELNCRTASCNSINWLRDLQVKGFEYFLNVSTRIWNGTFAASTPQTQVLTAAEID 1090
Qy 1058 LFNDSVFTL-----LPQGGAFVRSQTETKVEPEF-VENPLPLIVGSSVGLLLALITA 1110
Db 1091 TYNPQIYVIBENTVIP-----LTINKPEKVEVPTGVIVGSVIAGILLALLVA 1140
Qy 1111 ALYKLGPFKRQYKDM 1125
Db 1141 ILWKLGFPRKRYKRM 1155

RESULT 12
ITR2_MOUSE
ID ITR2_MOUSE STANDARD; PRT; 1178 AA.
AC Q62469; Q62163;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-2 precursor (Platelet membrane glycoprotein Ia) (GPIa)
DE (Collagen receptor) (VLA-2 alpha chain) (CD49b).
GN ITR2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_TaxID=10090;
RP SEQUENCE FROM N.A.
RSTRAIN=C57BL/6 X CBA; TISSUE=Lung;
```

```
RX MEDLINE=94363406; PubMed=8081889;
RA Edelman J.M., Chan B.M., Uniyal S., Onodera H., Wang D.Z.,
RA Damjanovich L., Latzer D.B., Pinberg R.W., Bergelson J.M.;
RT "The mouse VLA-2 homologue supports collagen and laminin adhesion but
RT not virus binding.";
RL Cell Adhes. Commun. 2:131-143(1994).
RN [2]
RP SEQUENCE OF 450-1178 FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=94355691; PubMed=7521231;
RA Wu J.B., Santoro S.A.;
RT "Complex patterns of expression suggest extensive roles for the alpha
RT beta 1 integrin in murine development.";
RL Dev. Dyn. 199:292-314(1994).
CC -I- FUNCTION: INTEGRIN ALPHA 2/BETA-1 IS A COLLAGEN RECEPTOR, BEING
CC RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER CELLS TO
CC COLLAGENS, MODULATION OF COLLAGEN AND COLLAGENASE GENE EXPRESSION,
CC FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED
CC EXTRACELLULAR MATRIX. IT IS ALSO A RECEPTOR FOR LAMININS, COLLAGEN
CC C-PROPEPTIDES AND E-CADHERIN. MICE HOMOZYGOUS FOR A NULL MUTATION
CC IN THE ALPHA-2 DIE VERY EARLY IN EMBRYOGENESIS.
CC -I- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-2
CC ASSOCIATES WITH BETA-1.
CC -I- SUBCELLULAR LOCATION: Type I membrane protein.
CC -I- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -I- SIMILARITY: Belongs to the integrin alpha chain family.
CC -I- SIMILARITY: Contains 1 VWFA domain.
CC -I- SIMILARITY: Contains 7 FG-GAP repeats.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; Z29987; CAA82877.1; --
DR EMBL; X75427; CAA53178.1; --
DR PIR; S44142; S44142.
DR HSP; P17301; IAOX.
DR MGD; MGI:96600; Itga2.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP_3.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWFA; 1.
DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
DR Platelet; Signal; Repeat; Calcium; Magnesium.
DR SIGNAL 1 26
DR CHAIN 27 1178 INTEGRIN ALPHA-2.
DR DOMAIN 27 1129 EXTRACELLULAR (POTENTIAL).
DR TRANSMEM 1130 1151 POTENTIAL.
DR DOMAIN 1152 1178 CYTOPLASMIC (POTENTIAL).
DR REPEAT 42 100 FG-GAP 1.
DR REPEAT ? 2 FG-GAP 2.
DR REPEAT 185 375 VWFA.
DR DOMAIN ? 483 FG-GAP 3.
DR REPEAT 431 483 FG-GAP 4.
DR REPEAT 485 546 FG-GAP 5.
DR REPEAT 548 607 FG-GAP 6.
DR REPEAT 612 664 FG-GAP 7.
DR REPEAT 496 504 POTENTIAL.
DR CA_BIND 560 568 POTENTIAL.
DR CA_BIND 624 632 POTENTIAL.
DR SITE 480 482 CELL ATTACHMENT SITE (POTENTIAL).
DR SITE 1154 1158 GPFR MOTIF.
DR DISULFID 80 89 BY SIMILARITY.
```

```

FT DISULFID 677 734 BY SIMILARITY.
FT DISULFID 786 792 BY SIMILARITY.
FT DISULFID 862 873 BY SIMILARITY.
FT DISULFID 1016 1047 BY SIMILARITY.
FT DISULFID 1052 1057 BY SIMILARITY.
FT CARBOHYD 102 102 N-LINKED (GLNAC. .) (POTENTIAL).
FT CARBOHYD 109 109 N-LINKED (GLNAC. .) (POTENTIAL).
FT CARBOHYD 429 429 N-LINKED (GLNAC. .) (POTENTIAL).
FT CARBOHYD 457 457 N-LINKED (GLNAC. .) (POTENTIAL).
FT CARBOHYD 472 472 N-LINKED (GLNAC. .) (POTENTIAL).
FT CARBOHYD 696 696 N-LINKED (GLNAC. .) (POTENTIAL).
FT CARBOHYD 1054 1054 N-LINKED (GLNAC. .) (POTENTIAL).
FT CARBOHYD 1071 1071 N-LINKED (GLNAC. .) (POTENTIAL).
FT CARBOHYD 1078 1078 N-LINKED (GLNAC. .) (POTENTIAL).
SQ SEQUENCE 1178 AA; 128926 MW; 1F194B9C0240F465 CRC64;

Query Match 18.0%; Score 1057; DB 1; Length 1178;
Best Local Similarity 27.8%; Pred. No. 3-2e-63;
Matches 341; Conservative 209; Mismatches 487; Indels 188; Gaps 45;

QY 1 FNLDTENAMTFQ--ENARFGQGVVQL---QGSRRVVGAPQEIIVAAQNGSLYOC--DYST 54
Dy : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
27 YNVLPGAKIFSGPSSQFGYSVQQLTNPQGNLLNGSPMSGPPENRMDGVYKCPVDLPT 86
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
55 GSCEPIRLQ-----VPVEAVNMSGLSLAAATSPQLLACGPTVHQTCSENTYKGLC 107
Dy : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
87 ATCEKMLQNSASISNTEIKTNMSLGLTLTRAPGTGGLTCGLPLMAHQCGNQYATGIC 146
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
108 FLFGSLNRLOOFPQ---KPEALRGCPQSDSDIAFLIDGSGSIIIPDFFRMKELVSTIMBOL 164
Dy : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
147 ----SDVSPDFQFLTSPFAVQACPSL--VDVVVVCDSESNISYP--WEAVKNFLVKFVTGL 199
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
165 K--KSKTLFSLMOYSEBFIHPTFKFQNNENPSRLIKPITQLLG-RTHATGLRKVVRE 221
Dy : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
200 DIGPKTKOVALIOYANEPRIIFINLNDPFTKEDVMQATSETROGGLDNTNTPRAIEFARDY 259
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
222 LFNTTNGARKNAFKILLLDGERFGDPLGVEDYVPELDRGVIRY---VLGF--GDAER 276
Dy : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
260 AYSQTSQGRPGATKVMVVMVVDGESH--DGSKLTKVIQCNDDIILRFGIAGVLGNRNALD 318
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
277 SEKSRQELNTVASPRDHFVQANFPAKTVQNLBEKIFATEGTQTSSSSPEHEMSQ 336
Dy : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
319 TKNLIKIKAIASPTERYFPFNVADEALKEKAGTLGEOIFSBIGTVQG--GDMFQMEMAQ 377
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
337 EGPSA---AITSNGPLLSVGSYDAGGVFLYTSKEKSTFTNMT--RVDSDMN--DAYLGYA 391
Dy : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
378 VGFSADYAPQNDIILMGNAGFADSGTLVQETSHKPVIFPKQAFDQVLQDNHSSFLGYS 437
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
392 AAILNRNVOSLVLGAPRYQHIGLVAMFRONTGMWESNANY----KGTQIGAYFGASLCS 447
Dy : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
438 VAAISTEDGVHFAVAGAFRANTGQIVLYSVNK---QGNVTVIQSHRGDQIGSYFGSLCS 494
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
448 VDVDSNGSTDVLVLCAGPHYBQTR--GGQSVCPPLPRQARWQCDAVLYGEOQPWGRF 505
Dy : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
495 VDVKDITIDVLLVGAPTYMNDLKSEKGVLFITKIGILNQHO---FLEGPEGTGNARP 551
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
506 GAALTVLDVNGDKLTDVAIGAPGEEDNRGAVYLFHGTSGSGISPSQSRIAGSKLSPR- 564
Dy : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
552 GSAIAALSDINMDGFNDVIGSPVENESGAVIYNGHQGT--IRTKYSQKILGNSGAFRR 610
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
565 -LOVEGSLSCGQDLTWGLDGLTVGAGHVLRLRSQPLVRLVKALMEFPNREVARNYPEC 623
Dy : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
611 HLQFFGRSLQGLNGDSITDVSIGALGVQIQLWSQSIADVALEAFTP----- 660
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
624 NDQVVKGEAGEVRVCLHVQKSTDRLRREGQIQSVVTVYDIALD----SGRPHSAVFNET 679
Dy : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
661 -DKITLLNKDAKITLKLCFRAEPFAGQNNQV--AILFNMTLDADGHSSRVTSRGVFRN 717
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
680 KNSRTRQTVGLHQTCTET--LKQLPNCIEDPSPVIRLNFSLVGTPTPSAFGNLBPVL 737
Dy : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
718 SERFLQKNVYVNEQVKCEHHSIQKPS---DVVNPLDLRVDISLENPGTS-----PAL 768
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
738 ABDAQRLFTLFPPEKKNCGNDNIQQDDLSI-----TPSFMSLDCLVVGGPREFNVT 788

```


Db 90 ATCEKLNQSTSPINVTETKTNISGLILTRNMTGGSLTCGLPLWAQCCNQYTTGVC 149
QY 108 FLFGSNLRQPOKPFALRGCCQEDSDIAFLIDGGSIIPHPFRMKELVSTIMEQLK-- 165
Db 150 SDISDPF-QLSASFSPATOPCSL-IDVVVCDENSIYPWD--AVXNPLEKFVQGLDIG 205
QY 166 KXKTLPSLMQYSEFRHIHTFKBPQNNPRSLIKPITOLLG-RHTATGLRKRVVRELFN 224
Db 206 PTKTQGLIYANNRPPVFNLTNTYKTBEMIVATSIQTSQYGGDLNTFGAIQYARKIAYS 265
QY 225 ITNGARKNAFKILFLITDCEKEDPLUGYEDVTPELDREGVIRY---VLGF--GDAPRSEK 279
Db 266 AASGGRSRTKVMVVVDGESH-DGSMKAVIDOCNHLIFAGIYAVLYLNALDYFN 324
QY 280 SRQELNTVASKPPDRHVQFNFAKTPQNLREKIPAFIEGTQGTGSSSSFEHNSQGF 339
Db 325 LIKEIKAIASIPTRYFFNVSDAALLKAGTLGEQIFSIETGVQ-GDNFQKMSQVGF 383
QY 340 SAALTSNGP--LLSTVGSYDAGGVPLYTSKEKSTFINMT--RVSDNM-DAYLVGAAR 394
Db 384 SADYSQNDILMLGAVGAFGSGTIVQKTSKHLIFPKQAFQDQILQDRNHSSTLIGSVAA 443
QY 395 ILENRVQSLVLAGPRYCHILGLVAMFRONTGMESNANV-----KGTQIGAFYGASLCSV 448
Db 444 ISTGESSTHVAGAPRANYTGQVLYSVN-----ENGNIPTVIAHRDQIGSYFGSVLCV 498
QY 449 DVDSNGSTDLVIGABHYEOTR--GGQVSVCLPRQAPRQOCDAVLGEQOGPWGRG 506
Db 499 DVDKDTITDVLVAGAPMYSDLCKEGRVLYFTIKKILGQHQ---FLEGPGIENTRFG 555
QY 507 AALTVDGVNKGDKLTDVAICAPEEDNRGAVLYFHGTSGSGISPSHSQRIAGS--KLSPR 564
Db 556 SAIAALSDINDMGFNDVIVGSPLENQNSGAVIYNHGQT-IRTKYSQILSDGAFRSH 614
QY 565 LQYFGOSLSGGQDLWDGLVGLVGAQGHVLLRSQPLVRKVAIMEFNPVARNVFEON 624
Db 615 LQYFGRSLDYGDLNGSDITDSIGAFQGVQVLWSQSIADVAIEASFTEKI--TLVKN 672
QY 625 DOYVKGKEAGEVRVCLHVQKSTFDRREGIQSVWVYDLALD----SGRPHSAFVNETK 680
Db 673 AQII-----LKCF-----SAKFRTKQNNQVAIVINITLDAGFSRVTSGCLFKENN 721
QY 681 NSTRRQTQVLGLTQC--ETLKLQLPNCIEDPVSPIVLNLFSLVGTPLSAPGNLRPLVA 738
Db 722 ERLQKNVNVQAQSCPEHIYIOEPS---DVVNSLDLRVDSLENPGTS-----PALE 772
QY 739 EDQRLFTALPPEKNGNDNICQDLSITF-----SFMSLDCLVVGCPREFFVTVTRND 794
Db 773 AVSETAKVFSIPPHKDCGDLGCLSDLVLDVROI PAAQEQPFIIVSNQNRILFTSVLKXK 832
QY 795 GEDSYRTQVTFPPFLDLVYKVKSTLQNRQSQRSLWRLACESAST-EVSGALKSTSCSINH 853
Db 833 RESAYNTGIVVDFSENLF-----ASFSLPVDGTEVTCQVAASQKSQVACDVCY 880
QY 854 PIPENSEVTFTNTPVDVSKASIGNKLLKANVTSENMPRTNKTFQLELPVKYAVYV 913
Db 881 PALKREQQVTFITFDFNLQ-NLQNOASLSFQALSESQENKADNLVNLKILPLLYDAEI- 938
QY 914 VTSHGVTKYLNFTASENTSRVMOHOVQVSNLQOR-----SLPISLVFLV----- 958
Db 939 ---HLRSTNINEYFESSDGNVPSIVHSEFVCPKFIPLSKVTGTSVPVSMATVHIHIQ 995
QY 959 -----PVRNLQNTVMDRPOVTF-SENLSSTCTHKEK 988
Db 996 YTKENKPLMYLTQVTDKAGDISCNADINPLKIGQT-----SSSVSPKSENF- -HTKE- 1047
QY 989 LPSSDPLAELRKAPVNVCSIAVCORIQCDIPFGIQEEFNATLKNLSFDWIKYKTSNH 1048
Db 1048 -----LNCRTASCSNVTWIKLVHMKGEFFVNVVTRINNGFASSTFQT 1091
QY 1049 LLIVSTAEI-LFNDVSFTLLPGCAFVRSQETKTEFFVPPNP-----LP--LIVGSSVG 1100
Db 1092 VQLTAAAEINTYNPETVI-----EDNTVTIPLMINKPDEAEVPTGTIIGSIIA 1141

QY 1101 GLLLLALITALYALYKLGPFKQYKDM 1125
Db 1142 GILLALLVALVILKLGPFKRYKYM 1166

RESULT 14

ITAL RAT
ID ITAL RAT STANDARD; PRT: 1180 AA.
AC P18614;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Integrin alpha-1 precursor (Laminin and collagen receptor) (VLA-1)
DE (CD49a).
GN ITGAL.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90338125; PubMed=2380249;
RA Ignatius M.J., Large T.H., Houde M., Tawil J.W., Barton A.,
RA Esch P., Carbonetto S., Reichardt L.P.;
RT "Molecular cloning of the rat integrin alpha 1-subunit: a receptor
for laminin and collagen";
RL J. Cell Biol. 111:709-720(1990).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 151-364.
RX MEDLINE=99313197; PubMed=10386626;
RA Nolte M., Pepinsky R.B., Vanyaminov S.Y., Koteliarsky V.,
RA Gotwals P.J., Karpusas M.;
RT "Crystal structure of the alphabeta1 integrin I-domain: insights into
integrin I-domain function";
RL FEBS Lett. 452:379-385(1999).
CC -!- FUNCTION: INTEGRIN ALPHA-1/BETA-1 IS A RECEPTOR FOR LAMININ AND
COLLAGEN. IT RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-F-P-G-
E-R IN COLLAGEN.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-1
ASSOCIATES WITH BETA-1.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X52140; CAA36384.1; -.
DR PIR; A35854; A35854.
DR PDB; 1CK4; 03-MAY-00.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR SMART; SM00191; int_alpha; 5.
DR SMART; SM00327; vwa; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWFA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
FT SIGNAL 1 28
FT CHAIN 29 1180 INTEGRIN ALPHA-1.
FT DOMAIN 29 1142 EXTRACELLULAR (POTENTIAL).

[illegible]

Db 602 ASMPHALSYFGSRVDRGLDLDGDDLDVAVGAQGAAILSSRPVHLTPSLVTPQAI SV 661
 Qy 619 NVEECNIOVVKGEAG--EYEVCLHVOKSTEDRLREGQIQSVVYTDALDSGRPHRAVF 676
 Db 662 VQDCRR---RGQAVCLTALCFQVTSRTPGRWDH---QFYMRFTASLDEWTAGARAAF 715
 Qy 677 NET--XNSTRQTOVLGHTQCTETKLQFNCEIDPVPSPVLRNLFSLVGTPLSAFGLNR 734
 Db 716 DSGQRLSPRRRLSVG-NVCEQLHPTHVLD-TSDYLRPVALTPTPALDNTTRPG----- 768
 Qy 735 PVLAEDAQRULTALFPFEKQCNENICQDDLSITFSFMSLDC-----LVVGGPREFN 786
 Db 769 PVLNEGSPTSIQKLVPPSKGPGNECVTDVLQ---VNMDIRGSRKAPFVVRGGRKVL 825
 Qy 787 VTVVRNDGSDSVYQVTFPFFPLDLSYKYSTLQNSORSWRLACESASSTEVSGALKS 846
 Db 826 VSTTLNKNENAYNTSLIIFSRNL---HLASLTPQR-ESPIKVECAAPSA-----HA 874
 Qy 847 TSCSINHPIFPENSEVTNITFDVDSKASIG---NKLL-----LKANVTSNNMPTNKT 898
 Db 875 RLCSVGHVFTQGAQVTFLLFEFEFSCSLLSQVFGKLTASSDSLSLERNGTLOENTAOQ--- 931
 Qy 899 EFQLELPVKVAVVNVVTSVSHGVSTKYLNFNTASENTSRVMOHOY-----QVS 943
 Db 932 -----SAYIQVEPH-----LLFSSESTLHYEVHPYGTLPVGPQPEKFTILRVQ 975
 Qy 944 NLG---QRLPISLVFLVP-----VRINQTVIWRDPQVTFSENLSSTCHTKERLPS 991
 Db 976 NLGCYVVSGLIISA--LLPAVAGGNYFLSLSQVI-----TNNASCIVQNLTEPPG 1024
 Qy 992 HSDPLAELRKAPVVCNSIAVCQRIQCDIPFFGIQEEFNATLKGNLSPDWIKTSHNHLI 1051
 Db 1025 PPVHPELQHTNRLNGSNTQCVVRCHLQGLAKGTEVSGVGLLRVHNEFFRRAKFKSLTV 1084
 Qy 1052 VSTABILFNDSVFTLLPGQGAFVRISQVETKVPPEVNPPLPLIVGSSVGGLLILALITAA 1111
 Db 1085 VSTFELGTEEGSVLQLTEASRWSESLLEV-VQTRPILISLWIMIGSVLGGLLMALLVFC 1143
 Qy 1112 LYKLGPF-----KROYK 1123
 Db 1144 LMKLGPPAHKKIPEEKREK 1164

Search completed: June 7, 2004, 17:12:42
 Job time : 13.9719 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OK protein - protein search, using sw model

Run on: June 7, 2004, 16:56:30 ; Search time 39.4889 Seconds
(without alignments)
9084.693 Million cell updates/sec

Title: US-09-902-481b-3
Perfect score: 5879
Sequence: 1 FNLDYENAMTFQENARGFGQ.....PKKQYKDMSEGGPPGAEPPQ 1137

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_25:
1: sp_archaea:
2: sp_bacteria:
3: sp_fungi:
4: sp_human:
5: sp_invertebrate:
6: sp_mammal:
7: sp_mhc:
8: sp_organelle:
9: sp_phase:
10: sp_plant:
11: sp_rodent:
12: sp_virus:
13: sp_vertebrate:
14: sp_unclassified:
15: sp_rviri:
16: sp_bacteriap:
17: sp_archaeap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4358	74.1	1151	11 Q9J130	Q9j130 rattus norv
2	3878.5	66.0	1036	11 Q8CA73	Q8ca73 mus musculu
3	3786.5	64.4	920	6 Q28984	Q28984 sus scrofa
4	3457	58.8	1169	4 Q81VA6	Q81va6 homo sapien
5	3297.5	56.1	1169	11 Q9QXH4	Q9qxh4 mus musculu
6	3225.5	54.9	1161	11 Q9QVE7	Q9qve7 rattus norv
7	1512.5	25.7	1161	11 Q9MTV4	Q9mtv4 mus musculu
8	1502	25.5	1160	11 Q9K200	Q9k200 mus musculu
9	1382	23.5	1196	13 Q9STF1	Q9stf1 cyprinus ca
10	1341.5	22.8	1086	4 Q96HB1	Q96hb1 homo sapien
11	1329.5	22.6	1187	13 Q98TF0	Q98tf0 cyprinus ca
12	1264	21.5	927	6 Q8HZV0	Q8hzv0 bos taurus
13	1142.5	19.4	1167	11 Q88340	Q88340 rattus norv
14	1097	18.7	1167	11 Q88341	Q88341 rattus norv
15	1038.5	17.7	1188	11 Q7TQC3	Q7tcq3 mus musculu
16	1034	17.6	1171	13 Q42094	Q42094 gallus gall

17	1020	17.3	1038	11 Q8BS01	Q8bs01 mus musculu
18	993	16.9	895	11 Q9WUF8	Q9wuf8 mus sp. itg
19	986.5	16.8	1160	6 Q8MKF4	Q8mkf4 felis silve
20	865	14.7	348	4 Q8TES5	Q8tes5 homo sapien
21	837	14.2	1332	5 Q9BPQ8	Q9bpq8 halocynthia
22	784	13.3	205	11 Q63001	Q63001 rattus norv
23	753.5	12.8	780	13 Q06271	Q06271 xenopus lae
24	738	12.6	823	4 Q8WY18	Q8wy18 homo sapien
25	686.5	11.7	823	11 Q8CE84	Q8ce84 mus musculu
26	672	11.4	1032	11 Q61989	Q61989 mus musculu
27	643	10.9	1036	11 Q91YD5	Q91yd5 mus musculu
28	639.5	10.9	1033	6 Q9BGU3	Q9bgu3 bos taurus
29	628.5	10.7	257	11 Q8C270	Q8c270 mus musculu
30	625.5	10.6	1474	5 Q86G87	Q86g87 pseudoplusi
31	619.5	10.5	1041	5 Q9UB90	Q9ub90 lytechinus
32	614.5	10.5	1041	5 Q76378	Q76378 lytechinus
33	579.5	9.9	1054	5 Q9U6S1	Q9ue61 strongyloe
34	579	9.8	1034	13 Q98TF7	Q98tf7 gallus gall
35	555.5	9.4	1053	11 Q80YP5	Q80yp5 mus musculu
36	546	9.3	1033	13 Q42598	Q42598 xenopus lae
37	546	9.3	1036	6 Q7YRP8	Q7ytp8 equus cabal
38	534	9.1	1016	13 Q91779	Q91779 xenopus lae
39	530	9.0	1047	6 Q9WZD6	Q9wzd6 bos taurus
40	529	9.0	974	11 Q924W2	Q924w2 rattus norv
41	529	9.0	1007	6 Q9GK48	Q9gk48 bos taurus
42	528	9.0	1073	11 Q8CC06	Q8cc06 mus musculu
43	526.5	9.0	1119	5 Q86G88	Q86g88 pseudoplusi
44	521.5	8.9	1132	11 Q80Z18	Q80zi8 mus musculu
45	513.5	8.7	1034	6 Q9TUN4	Q9tun4 oryctolagus

ALIGNMENTS

RESULT 1

Q9J130 PRELIMINARY; PRT; 1151 AA.
AC Q9J130
DT 01-OCT-2000 (TREMELrel. 15, Created)
DT 01-OCT-2000 (TREMELrel. 15, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Integrin beta 2 alpha subunit.
OS Rattus norvegicus (Rat)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Fathallah D.M. Sr., Zerriah K. Jr.;
RT Cloning of the rat CD11b cDNA sequence."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF268593; AAF81280.1; -
DR HSSP; P11215; 1BHQ.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; P:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00342; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWFA; 1.
SQ SEQUENCE 1151 AA; 126943 MW; 8F785695D4074CA5 CRC64;

Query Match 74.1%; Score 4358; DB 11; Length 1151;

Best Local Similarity 72.3%; Pred. No. 1.2e-314; Indels 2; Gaps 2;
Matches 822; Conservative 151; Mismatches 162;

QY 1 FNLDTENMTFOENARGFQSVVLOGSRVYVGAPOEIVAAHQSGSLYOCYSTGSCPEI 60
DB 17 FNLDTENMTFOENARGFQSVVLOGSRVYVGAPOEIVAAHQSGSLYOCYSTGSCPEI 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPPQLACGPTVHQTCSENYVYKGLCFGLGSLNRQPOK 120
DB 77 PLQVPVEAVNMSLGLSLAATTSPPQLACGPTVHQTCSENYVYKGLCFGLGSLNRQPOK 136
QY 121 FPEALRGCCQEDSDIAFLDGSISITPHDRFMKELVSTIMEQLKSKTLPSLMQYSEEF 180
DB 137 FPEALRGCCQEDSDIAFLDGSISITPHDRFMKELVSTIMEQLKSKTLPSLMQYSEEF 196
QY 181 RIHFTFKFQNNPNSLKIPIKIPITQLLGRTHATGLRQVRELFNITNGARKNAFKILFL 240
DB 197 RIHFTFNDPKNDPKSHVPEIRQLNGRTKTASGIRKVVRELFOKINGARDNAAKILVVI 256
QY 241 TDEKFGDGLYEDVITPELDREGVIRYVGLGDAFRSEKSRQELNVTASKPRDHVFOAN 300
DB 257 TDEKFGDGLYEDVITPELDREGVIRYVGLGDAFRSEKSRQELNVTASKPRDHVFOAN 316
QY 301 NPEALKTQVQUREKIPIKIPITQLLGRTHATGLRQVRELFNITNGARKNAFKILFL 360
DB 317 NPEALKTQVQUREKIPIKIPITQLLGRTHATGLRQVRELFNITNGARKNAFKILFL 376
QY 361 GVELYTSKEKSTFINTRVDSMDNDAYLGYAAAIILNRVQSVLGLAPRYCHIGLVAMFR 420
DB 377 GVELYTSKEKSTFINTRVDSMDNDAYLGYAAAIILNRVQSVLGLAPRYCHIGLVAMFR 436
QY 421 QNTGMESNANVAGTIGAYFGASLSDVDNSGSDLVILGAPHYVYOTRGQVSVCP 480
DB 437 QNTGMESNANVAGTIGAYFGASLSDVDNSGSDLVILGAPHYVYOTRGQVSVCP 496
QY 481 PRQGRARWQDAVLGYEGQOPWGRFGAALTFLGDNVNGDKLTDAI GAPBEENRGAVYLF 540
DB 497 PRQGRARWQDAVLGYEGQOPWGRFGAALTFLGDNVNGDKLTDAI GAPBEENRGAVYLF 555
QY 541 HGTSGSGISPSHQRTAGSKLSRLQYQSGSLGGDLTMDGLVTLVGAQGHVILLRQ 600
DB 556 HGTSGSGISPSHQRTAGSKLSRLQYQSGSLGGDLTMDGLVTLVGAQGHVILLRQ 615
QY 601 PVLVRKAIMEFNPREVARNYFECNDQVVKGEAGEVRVCLHVOKSTRDLRLEGQIOSVT 660
DB 616 PVLVRKAIMEFNPREVARNYFECNDQVVKGEAGEVRVCLHVOKSTRDLRLEGQIOSVT 675
QY 661 YDLALDSGRPHSAVNETKSTRTQVTLGLTCTETLKLQLPNCIEPVSPIVLRNLF 720
DB 676 YDLALDSGRPHSAVNETKSTRTQVTLGLTCTETLKLQLPNCIEPVSPIVLRNLF 735
QY 721 SLVGTPLSAFQNLPRVLAKDAQBLTALPFFKNCNGNDNICQDDLSITTSFMSLCLVVG 780
DB 736 SLVGTPLSAFQNLPRVLAKDAQBLTALPFFKNCNGNDNICQDDLSITTSFMSLCLVVG 795
QY 781 GPREFNVTVRNDGDSYQVTPFPPLDLSTRKUSTLQNRQSRWLACASSTEV 840
DB 796 GPREFNVTVRNDGDSYQVTPFPPLDLSTRKUSTLQNRQSRWLACASSTEV 854
QY 841 SGALKSTSCSINPIIPENSEVFNITFDVDSKASLGNKLLKANTVSENMPRTNKTEF 900
DB 855 SGALKSTSCSINPIIPENSEVFNITFDVDSKASLGNKLLKANTVSENMPRTNKTEF 914
QY 901 QLELPVAYIYVMTSGVSTKYLINFASNTGRVMOHQVQVGNLQORSPLISLVLV 960
DB 915 QLELPVAYIYVMTSGVSTKYLINFASNTGRVMOHQVQVGNLQORSPLISLVLV 974
QY 961 RLQVTVLWDRPOVTFSENLSSTCHTKERLPSHSDLAELARKAPVNCISIAVOCIOCDIP 1020
DB 975 RLQVTVLWDRPOVTFSENLSSTCHTKERLPSHSDLAELARKAPVNCISIAVOCIOCDIP 1034
QY 1021 FFGIQEBFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQAFVRSQTET 1080
DB 1035 FFGIQEBFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQAFVRSQTET 1094
QY 1081 KVEFFEPVNPPLVGVSSVGLLILLALITAGLYKLGFFKQYKDMNSEGGPPGAPQ 1137

DB 1095 KVEFFEPVNPPLVGVSSVGLLILLALITAGLYKLGFFKQYKDMNSEGGPPGAPQ 1151
RESULT 2
Q8CA73 PRELIMINARY; PRT: 1036 AA.
AC Q8CA73;
DT 01-MAR-2003 (TREMREL. 23, Created)
DT 01-MAR-2003 (TREMREL. 23, Last sequence update)
DT 01-OCT-2003 (TREMREL. 25, Last annotation update)
DE Integrin alpha M.
GN ITGAM OR F730045J24RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spinal cord;
RX MEDLINE=22354683; PubMed=12466851;
RA THE FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
DR SMEL; AK039444; BAC30350.1; -.
DR PIR; PT0572; PT0633.
DR PIR; PT0697; PT0697.
DR MGD; MGI:96607; Itgam.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; C:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR Pfam; PF01839; FG-GAP; 1.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VMPA; 1.
SQ SEQUENCE 1036 AA; 115126 MW; 4F4F39BPE1E188E77 CRC64;
Query Match 66.0%; Score 3878.5; DB 11; Length 1036;
Best Local Similarity 65.6%; Pred. No. 4.7e-279;
Matches 747; Conservative 126; Mismatches 146; Indels 119; Gaps 2;
QY 1 FNLDTENMTFOENARGFQSVVLOGSRVYVGAPOEIVAAHQSGSLYOCYSTGSCPEI 60
DB 17 FNLDTEHPMTFOENAKFGQSVVLOGSGTGVVVAAPQEAQVAVNQTGALYCCDYSTRCHPI 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPPQLACGPTVHQTCSENYVYKGLCFGLGSLNRQPOK 120
DB 77 PLQVPVEAVNMSLGLSLAATTSPPQLACGPTVHQTCSENYVYKGLCFGLGSLNRQPOK 136
QY 121 FPEALRGCCQEDSDIAFLDGSISITPHDRFMKELVSTIMEQLKSKTLPSLMQYSEEF 180
DB 137 FPEALRGCCQEDSDIAFLDGSISITPHDRFMKELVSTIMEQLKSKTLPSLMQYSEEF 196
QY 181 RIHFTFKFQNNPNSLKIPIKIPITQLLGRTHATGLRQVRELFNITNGARKNAFKILFL 240
DB 197 RIHFTFNDPKNDPKSHVPEIRQLNGRTKTASGIRKVVRELFOKINGARDNAAKILVVI 256
QY 241 TDEKFGDGLYEDVITPELDREGVIRYVGLGDAFRSEKSRQELNVTASKPRDHVFOAN 300
DB 257 TDEKFGDGLYEDVITPELDREGVIRYVGLGDAFRSEKSRQELNVTASKPRDHVFOAN 316
QY 301 NPEALKTQVQUREKIPIKIPITQLLGRTHATGLRQVRELFNITNGARKNAFKILFL 360
DB 317 NPEALKTQVQUREKIPIKIPITQLLGRTHATGLRQVRELFNITNGARKNAFKILFL 376

```
QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAIILNRVQSLVLCAPRYQHIGLVAMFR 420
DB 377 GAFLYTSKDKVFINNTRVDSMDNDAYLGAAAIILNRVQSLVLCAPRYQHIGLVAMFR 436
QY 421 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLGAPHYEQTRGGQVSVCP 480
DB 437 ENFGTWEPTSISKG----- 450
QY 481 PRQARWQCDVAVLYGEOQPMGRFGAALTVLGVNGDKLTDVAIGAPGEEDNRGAVYLF 540
DB 451 ----- 450
QY 541 HGTSGSISPSHSORISAGSKLSPRLOYFGOSLGGDLTMDGLVDTLTVCAOCHVLLRSO 600
DB 451 -----SRIIGAHSFPGLOYFGOSLGGDLTMDGLVDTLTVCAOCHVLLRSO 498
QY 601 PVLVRKAIEMFNPREVARNVFCNDQVVGKAGEVRVCLHVQKSTRDLREGOIQSVYT 660
DB 499 PVLRLBATWFSFKKVARSVFACQEQVLKNDKAGEVRVCLVRKNTKORLREGDIQSTVT 558
QY 561 YDLALDSGPHSRVAVNETNSTRTOTVGLTQTCETILKLOIPNCIEDPUSPVLRLNP 720
DB 559 YDLALDPGSRIRAPDETNNTRTQVFLMQKCEITLKLILPCVDDSVSPTILRLNY 618
QY 721 SLVGTPLSAFGNLRPVLAEDAORLFTALPFFKKNCGNDNI CODDLSITPSFMSLCLVVG 780
DB 619 TLVGEPLRSFGNLRPVLAEDAORLFTALPFFKKNCGNDNI CODDLSITPSFMSLCLVVG 678
QY 781 GPREFNVTVVRNDGSDSVRTQVTFPPFLDLSYRKVSTI ONORSORSWEL-ACESASSTE 839
DB 679 GPDEFNNSVTLRNDGSDSVRTQVTFPPFLDLSYRKVSTI ONORSORSWEL-ACESASSTE 738
QY 840 VSGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANKVTSENNMPTNKE 899
DB 739 GHGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANKVTSENNMPTNKE 798
QY 900 FOLELPVKYAVYVTVSHGVSTKYLNFTASENTSRVMQHQYQVSNLQGRSLPISVLFP 959
DB 799 FOLELPVKYAVYVTVSHGVSTKYLNFTASENTSRVMQHQYQVSNLQGRSLPISVLFP 858
QY 960 VRLNQTVIWRDPQVTSSENLSTCHTKERLPSHSDFLAELKAPVNCVSIANCORICODI 1019
DB 859 VQINNVTVWHPQVIFSQLSSACHTEQSPHSNFRDQLETPVLNCSVAVCKRIQCOL 918
QY 1020 PFGIOEFNATLKGSLSPDWYIKTSHNLLIVSTAEILFNDSVFTLLPFGQAFVRSQTE 1079
DB 919 PSFNTQEIFNVTLKGNLSPDWYIKTSHGHLVLSSTELLFNDSAPALLPQSSVYRSKTE 978
QY 1080 TKVEPEVPNPDLPIVSSVGGLLILALITAALYKLGFFKQYKDMKSGGPPGAPQ 1137
DB 979 TKVEPEVHPNPDLPIVSSVGGLLILALITAGLYKLGFFKQYKDMKNEAAPQDAPPQ 1036

RESULT 3
QID Q28984 PRELIMINARY; PRT; 920 AA.
AC Q28984;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-FEB-1997 (TEMBLrel. 02, Last sequence update)
DT 01-JUN-2003 (TEMBLrel. 24, Last annotation update)
DE Cdlb (Fragment).
GN Cdlb.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sue.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Lee J.-K., Schok L.B., Rutherford M.S.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U40072; AAB16869.1; -.
DR HSSP; P11215; 1BHQ.
```

```
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; EG-GAP; 3.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS50234; VWFA; 1.
FT NON_TER 1
FT NON_TER 920
SQ SEQUENCE 920 AA; 102440 MW; E96CC51E350DD5AC CRC64;
```

Query Match 64.4%; Score 3786.5; DB 6; Length 920;
Best Local Similarity 78.4%; Pred. No. 2.7e-272;
Matches 722; Conservative 85; Mismatches 113; Indels 1; Gaps 1;

```
QY 118 POKPEALRGCPQSDSDIAFLIDGSGSIIPHDPRMKELVSTIMEQLKSKTFLSLMYS 177
DB 1 POKPEALRGCPQSDSDIAFLIDGSGSIINRLDFQMKEFVSTVMQFQKSKTFLPALMYS 60
QY 178 SEFRIHPTFKBFQNNPNSRLIKPITQLLGHRTATGLRKVVRHLEFNITNGARKNAFIL 237
DB 61 EDFYTHPTFNDPKZNPSPKLLVRPIQLLGRTHRTATGIRKVRHLEFHSKGARENALKIL 120
QY 238 FLLTDCGKFGPLGVEDVPELDRGVIRYVLGPDAPRSKRSQELMTVASKPPDRHVF 297
DB 121 VITDGEKFGPLGVEDVPEADRKGVIRYVGVGDAPNSKREELNTASKPCGDHVF 180
QY 298 QANNPEALKTVQNLREKIPAIETQTGSSSFHEMSQEGFSAAITNGPILLSTVGSYD 357
DB 181 QVNNFEAVKTIQNLQOEKTFAIETQTGSSSFCEMSQEGFSAAITNGPILLGAVGSPD 240
QY 358 WAGGVLYTSKESKTFINMTVDSDMDAYLGAAAIILNRVQSLVLCAPRYQHIGLV 417
DB 241 WAGGAFLEMPKDRVTFINTTRVDSMDNDAYLGAAAIILNRVQSLVLCAPRYQHIGLV 300
QY 418 MPRONTGHWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLGAPHYEQTRGGQVSV 477
DB 301 MFKQNSGAWENKADIKSGISQISYFGASLCSVDVNRDSSDLVLGAPHYEQTRGGQVSV 360
QY 478 CPLPRGQARVQCDVAVLYGEOQPMGRFGAALTVLGVNGDKLTDVAIGAPGEEDNRGAV 537
DB 361 CPLPGQ-RARQCRVILCGEQHPWSRFGNALTAGLVNGDKLTDVAIGAPGEEDNRGAV 419
QY 538 YLFHGTSGSISPSHSORISAGSKLSPRLOYFGOSLGGDLTMDGLVDTLTVCAOCHVLL 597
DB 420 YLFHGTSELGSPSHSORISAGSKLSPRLOYFGOSLGGDLTMDGLVDTLTVCAOCHVLL 479
QY 598 RSQPVLRVKAIEMFNPREVARNVFCNDQVVGKAGEVRVCLHVQKSTRDLREGOIQS 657
DB 480 RSQPVLRVEASNVFKPREVARNVYECRQAAKTQIAGEVQVCLQVRKSTWDLREGDQTS 539
QY 558 VVYDLDLDSGRPHSRVAVNETNSTRTOTVGLTQTCETILKLOIPNCIEDPUSPVLRL 717
DB 540 ITTYDLDLDPGRPHSRVAVNETNSTRTOTVGLSRKCEHLALWLPDVEDSVTFVILR 599
QY 718 LNFSLVGTPLSAFGNLRPVLAEDAORLFTALPFFKKNCGNDNI CODDLSITPSFMSLCL 777
DB 600 LNFSLVGKFPASSFGNLRPVLAEDAORLFTALPFFKKNCGNDNI CODDLSITPSFMSLCL 659
QY 778 VVGGRPFNVTVVRNDGSDSVRTQVTFPPFLDLSYRKVSTI ONORSORSWELACESASS 837
DB 660 VVGGRPDLKVLTVVRNDGSDSVRTQVTFPPFLDLSYRKVSTI ONORSORSWELACESASS 719
QY 838 TEVSGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANKVTSENNMPTNKE 897
DB 720 TEESTALKSTSCSINHPIPPENSEVTFNITFDVDPDAPLGYKLLKANKVTSENNMPTNKE 779
QY 898 TEFQLELPVKYAVYVTVSHGVSTKYLNFTASENTSRVMQHQYQVSNLQGRSLPISVLFP 957
```



```

780 TEFOLELVYAVVVVTSLEVTSTKYNFTASEKTRHVIHQYCFNNLGOKLPISVWFV 839
958 VVRLNQTIVWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAAPVNCISAVCORIQ 1017
840 VVRLNQTIVWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAAPVNCISAVCORIQ 899
1018 DIPPFGIOEBFNATLKNLSF 1038
900 DIPSGIOEBELKVTLKNLSF 920

RESULT 4
Q81VAG PRELIMINARY; PRT; 1169 AA.
AC Q81VAG;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Integrin, alpha X (Antigen CD11C (p150), alpha polypeptide).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RC SEQUENCE FROM N.A.
PC TISSUE=Blood;
RA Strausberg R.;
RL Submitted (Sep-2002) to the EMBL/GenBank/DBJ databases.
RE EMBL; BC038237; AAH38237.1; -.
DR GO; GO:0008305; C:intrin complex; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR00413; Integrin_alpha.
DR InterPro; IPR002035; WVF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWA; 1.
DR SEQUENCE 1169 AA; 128521 MW; 178484FEFC79EB6 CRC64;

Query Match
Best Local Similarity 60.78; Pred. No. 1.3e-247;
Matches 685; Conservative 143; Mismatches 295; Indels 6; Gaps 4;

Qy 1 FNLDENAMTPOENARFGSGVLOQSGRVVVGAPQIVAVANQSGSLYQCDYSGCEPI 60
Db 20 FNLDTELTAFRVDGSGVSVQYANVWVVGAPQKITAANQGTGGLYQCGYSTGACEPI 79
Qy 61 RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSNTYVKGCLFPGSNLRQOPQ 120
Db 80 GLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHHECGRNWYLTGLCFLLGPT--QLTQR 137
Qy 121 PPEALRCQBDSDIAFLDGSIIIPHDPRMKELVSTIMEQLKSKTLPFLMQYSEEF 180
Db 138 LPVSRQECPRQBDIVFLDGSISISSRNTATMNFVRAVISOQRPSTQFSLMQFSKF 197
Qy 181 RIHTEKFNPNPNSRIKIPITOLGRTHATGLRKVVRLEFNIWGAARKNAFKILFL 240
Db 198 QTHTFEFRKSNPLSLASVHQLQGTFTYATAIQNVHRLFASVGAARDAAKILVI 257
Qy 241 TDGKFGDPLGYEDVPELDRGVIRYVLGPDAPRSEKSRQELNVTASPPRDHVFQAN 300
Db 258 TDGKFGDPLGYEDVPELDRGVIRYVLGPDAPRSEKSRQELNVTASPPRDHVFQAN 317
Qy 301 NFEALKTVQNLREKIPAEIGTQSGSSPHEMSQSGFSAITSNGPLSLTSGSYDMAG 360
Db 318 DFDALKIQNLREKIPAEIGTQSGSSPHEMSQSGFSAITSNGPLSLTSGSYDMAG 377

```

```

361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAILRNVRQSLVIGAPRYQHIGLVAMPR 420
378 GAFLYPPNMPSTFINNSQENVDMDSYLGSYSELALMKGQVQSLVIGAPRYQHTGAVIPT 437
421 QNTGKMSNANVKGTCIGAYFGASICSVDVDSNGSTDLVIGAPHYETQRCQGVSVCP 480
438 QVSRWRKAEVGTGIGSYFGASICSVDVDSNGSTDLVIGAPHYETQRCQGVSVCP 497
481 PRQARMQCDVLYGEOQPMGRFGAALTIVLGDVNGDKLTDVAIGAGEEDNRGAVYLF 540
498 PRGWR-RWCDVLYGEOQPMGRFGAALTIVLGDVNGDKLTDVAIGAGEEDNRGAVYLF 556
541 HGTSGSGISPSHSQRIAGSKLSPRLOYRQOSLSGGQDLTMDGLVLTCAQCHVLLRSQ 600
557 HGVLGSPISPSHSQRIAGSKLSPRLOYRQOSLSGGQDLTMDGLVLTCAQCHVLLRSQ 616
601 PVLRYKAIEMENPREVARNEPCNDQVYVKGKAGEVRVCLHVOKSTRDLRLEQIQSVVT 660
617 PVLWGVSMQFIPABIPRSAFECEQVVEQTLVQSNICLYIDKSKNLLSGRDLQSSVT 676
661 YDLALDSGPHSRVAFNETKSTRQTVGLTQTCETIKLQPCIEDPVPVILRLNF 720
677 LDALDPGRLSPRATFQTKNLSLRVRLGKAHCENFNLLPSCVEDSVTPTILRLNF 736
721 SLVGTPLSAFGNLRPVLAEQAORLFTALPPPEKNCNDNICODDLSITPFSMLDCLVVG 780
737 TLVKGKLLAFNLRPMLAADAQRYFTASLPPEKNCGADHICODNLGISFSPGLKSLVVG 796
781 GPRFNVTVTVNDGEDSVTRQVTFPFLDLSYRKVSTLQKRSORSWELACESASTEV 840
797 SNLELNARVWVNDGEDSYGTVTFSSHAGLSYRVVABEQKQGLRSLHLTCDSPVVG-- 854
841 SGALKSTCSINHPIPPENSEVTFNITFDVDSKASLGKLLKANKVTSNNMPTNKTEF 900
855 SGTWSTSCRIINHLIFRGGAQITFLATDVSPKAVLGDRLLTANVSSNNVPTSKTTF 914
901 QLELPKVAIVVTVSHGVSTKYLNFTAS-ENTSRVMOHQYQVSNLQBSLPSILVFLVP 959
915 QLELPKVAIVVTVSHGVSTKYLNFTAS-ENTSRVMOHQYQVSNLQBSLPSILVFLVP 974
960 VRLNQTIVWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAAPVNCISAVCORIQ 1019
975 VRLNQTIVWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAAPVNCISAVCORIQ 1034
1020 PFGIOEBFNATLKNLSFDPWYIKTSHNHLIVSTABILFNDQSVFTLLPQGAFFVRSQTE 1079
1035 PSFSVQEBLDPFLKGNLSFGWVRQILQKKVSVSVVAEITPDTSVYSQLPGEAPMKAQTT 1094
1080 TKVEPEVPEVNPFLIVGSSVGLLALLALITAAALYKLGPEKROYKDMSE 1128
1095 TVLEKTKVENPFLIVGSSVGLLALLALITAAALYKLGPEKROYKDMSE 1143

RESULT 5
Q9QXH4 PRELIMINARY; PRT; 1169 AA.
AC Q9QXH4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Leukocyte adhesion glycoprotein p150.95 alpha integrin subunit.
GN ITGAX.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
CX NCBI_TaxID=10090;
RN [1]
RC SEQUENCE FROM N.A.
RA Huang X., Gorski K., Tong C., Rattis F.-M., Tseng S.-Y., Pardoll D.,
RT Teuchiya H.;
RT Isolation of Genes Selectively Expressed by Dendritic Cells.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
EMBL; AF211864; AAF23492.1; -.

```

```
DR HSSP; P11215; 1BHQ.
DR MGP; MGI:96609; Itgac.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; P:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR GO; GO:0007229; P:integrin-mediated signaling pathway; IEA.
DR InterPro; IPR00413; Integrin_alpha.
DR InterPro; IPR02035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3_A.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS50234; VWEA; 1.
KW Integrin.
SQ SEQUENCE 1169 AA; 129150 MW; C616412033C219A6 CRC64;

Query Match 56.1%; Score 3297.5; DB 11; Length 1169;
Best Local Similarity 56.2%; Pred. No. 9.4e-236;
Matches 641; Conservative 174; Mismatches 306; Indels 19; Gaps 7;

Qy 1 FNLDENAMTFOENARFGQSVQLOGSRVVGVAQPIVAAQNGSLYQCDYSTGSCPEI 60
Db 20 FNLDABKLTHFMDGABFGHSLVQYDSVWVVGAPKEIKATNQIGGLYKGYHTGNCPEI 79
Qy 61 RLQVPEAVNMVNLGLSLAATTPPOLACGPTVHOTCSENTYVKGCLFGLGSLNRQPOK 120
Db 80 SLQVPEAVNISLGLSLAATNPWILLACGPTVHTCRENIYLTGLCLFLSSSEKQS-QN 138
Qy 121 PPEALRGCPQSDIAFLIDSGSIIIPDFRMKELVSTIMEQLKSKTILFSLMQVSEEF 180
Db 139 PPTAQOECPCQDQIVELIDSGSISSTDFPKMLDFKAVMSQLQRPSTFSLMQPFDYF 198
Qy 181 RIHTEFEFONNPNRSILKPIITOLLGRTHATGLRKVVRLENIWNGARKNAFKILFL 240
Db 199 RVHFTFNFNISTPSFLSLGSRVQLRGITTYTASAKHVIITELFTTQSGAKQDATKVLVI 258
Qy 241 TDGEKFGDPLGYEIVIFELDREGVIRYVGLFGDAPRSEKSRQELNTVASPPRDHVFQAN 300
Db 259 TDGRKQGNLSYDVIPEAAEASIIIRVAIGVKAFYNEHSHQELKALASMPSEHYVFSVE 318
Qy 301 NPEALKTVQNLREKIPIAETGTQTSSSSEHENSQEGSAATNSGPNLSTVSGSDWAG 360
Db 319 NFDALKDIENOLKEKIPIAETGTETPSSTPELEMSQEGSAVFTPDGVLGAVGVSFWSG 378
Qy 361 GVFLYTKSEKSTFTNTRVDSMDMDAYLGVAALILNRVQSLVGLGAPRYQHIGLVAMPR 420
Db 379 GAFILPSNMRTFFINMSENEDEDAVLGYSTALAFKGVHSLILGAPRHQHTGKVIFT 438
Qy 421 QNTGMWESNANKTQIGAYPGASLCSDVDVDSNGSTDLVLIGAPHYYEOTRGGQSVCP 480
Db 439 QESRHWPKSEVRGTQIGSVFGASLCSDVMDRDSGLVIGVPHYHYEHTRGGQSVCPM 498
Qy 481 PRGORAWOCDVLYGQGGQPGWFGAALTVLGVNGDKLTVDAIGAPGEEDNRGAYTLF 540
Db 499 P-GVGRWHCTTLHGQGHGHPWFGAALTVLGVNGDSLADVAIGAPGEENRGAYIF 557
Qy 541 HGTSGSGISPHSQRIAGSKLSPRLQYFGQSLSGQDLFWDLVLTGVAQGHVLLRSQ 600
Db 558 HGASRQDIAPSPQRSIGASQIPSIQYFGQSLSGQDLTRDGLVDLAVGSKGRVLLRTR 617
Qy 601 PVLKVKALMEPNPEVARNPECNDOVVKGEAGEVEVCLHVQKSTDRLRREGIQSVVT 660
Db 618 PILRVSTPVHTPAEIRSRVSECEQVAPBOTLSDATVCLIHESPKTL--GDLRSTVT 675
Qy 661 YDLALDSGRPHSAVFVETKNSTRQTVLGLTQTCETLKLQPLNCIEDPVSVTLRLNF 720
Db 676 PDLALDHGLRLSTRAIFXETKTRALTRVKTILGLNKHCSVKLLLPACVEDSVTPITLRLNF 735
Qy 721 SLVGTPLSAFGNLRFPVLAEDAQRLFTALFPPEKNCNGNDNICQDDLSTLTFMSLDCLVWG 780

736 SLVGVPISLONLQPMPLAVDDQTYFTASLPFEKNGCAGDHICQDDLSVFGFPLKTLVWG 795
781 GPREFNVTVTVNDGSDSYRTQVTFPPFDLDSVRKYSTLQ-----NORSORSWR 829
796 SDLELAVDVTVSNDGSDSYGTTVTFEPVGLSFRVABGQVFLRKKEQDQWQERGQSHLH 855
830 LACESASSTEVSGALKSTSCSINHPIFFENSEVTFNITFDVDSKASIGNKLKLVNYS 889
856 LACD--STDFRSQGLWSTSCSRHVFIRGSGQMTFLVTFDVSFKAGLDELRLARVGSB 913
890 NNMRPNKTEFOLELPVKYAVVNVTVSHGVSTKYLNFTASE-NTSRVMOHQYQVSNLQOR 948
914 NNVPGETPKTTFOLELPVKYAVVTVMISSHDQFTKYNLFSTSEKTSVBERFQVNNLQOR 973
949 SLPISLVFLVPLVNLQTVIWDROPQVTFSENLSSTCTTKERLPSHSDFLAELRKAPVNC 1008
974 DVPVSINFVWPIELKGEAVW-TVWNSHPQNPFLTCVVRNLKPTQFDLLTHMQSPVLD 1032
1009 IAVCORIQCDIPFGIOBEFNATLKNLSDFDVMYIKTSHNHLIVSTAELFNDSTVTL 1068
1033 IADCLHLRCDIPSGILDDELTFILKGNLSFGWISQTLQKVKLLSEAEITFNTSVISQLP 1092
1069 GQGARVRSOTETKVEPPEVNPPLIVGSSVGLLILALITRALYKLGREFKROYKMMSE 1128
1093 GQAFLEAQTKTVLEKYKVNVPVPLIVGSSVGLLILALITLTKAGFFKROYKEMLEE 1152

RESULT 6
Q9QY87 PRELIMINARY; PRT; 1161 AA.
AC Q9QY87;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Alpha D integrin.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA O'Brien M.M., VanderVieren M., Kilgannon P.D., Dietech G.,
RA Gallatin W.M.;
RT "Cloning of rat alpha D, a novel beta 2 integrin.";
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF021334; AAF21241.1; -.
DR HSSP; P11215; 1BHQ.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; P:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR GO; GO:0007229; P:integrin-mediated signaling pathway; IEA.
DR InterPro; IPR00413; Integrin_alpha.
DR InterPro; IPR02035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS50234; VWEA; 1.
KW Integrin.
SQ SEQUENCE 1161 AA; 126600 MW; 2258491A984A705E CRC64;

Query Match 54.9%; Score 3225.5; DB 11; Length 1161;
Best Local Similarity 57.2%; Pred. No. 2.1e-230;
Matches 645; Conservative 163; Mismatches 306; Indels 13; Gaps 9;

Qy 2 NLDTENAMTFOENARFGQSVQLOGSRVVGVAQPIVAAQNGSLYQCDYSTGSCPEIR 61
```

Db	21	NLDVEEPIVFREDAASFGQTVWFGSGRLVVGAPLEAVAVNQTRGLYDCCAPATGMCQPIV	80
Qy	62	LOVPVEAVNMSLGLSLAATTSPPOLLACGPTVHTCSENTYVKGCLFLGSLNRQQPKF	121
Db	81	LRSPLEAVNMSLGLSLATNNAQLACGPAQACVKNMYAGSCLLLGSSLL-QRIQAV	139
Qy	122	PEARLGCQPDSDIAFLIDSGSIIIPDPRMRELSTIMEOLKSKSLFSLMOYSEEF	181
Db	140	PASMECPROEMDIAFLIDSGSINQDFAQMDKDFKALMGEBFASSTLFLSLMOYNILK	199
Qy	182	IHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVVRELFNITNGARKNAKILFLLT	241
Db	200	THFTFTFBNKILDPQSLVDPTVOLQGLTYATGRTVMEELFHSKNSRKSANKILLVIT	259
Qy	242	DGEXFGDPLGYEDVIPELDREGVRYVVLGFGDAFRSEKSRQELNTVASKPPRDRHVFQANN	301
Db	260	DGQKRPDLEYSVDIPADKAGIIRYALGVGDARQDEPTALKELNTIGSAPPQDHPFRVGN	319
Qy	302	FEALKTVQNLUREKIFAIEGTQSGSSSFEHMSQEGFSAAITSNGLLTVGSDYDAGG	361
Db	320	FAALRSIQRLQEKIFAIEGTQSSSSSFFHEMSQEGFSALTSDGVLGAVGFSWSGG	379
Qy	362	VFLATSEKSTFFINMTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYOHIGLVAMFRQ	421
Db	380	APLPPTNTRPTFINMSQENVDRSYLGYSTAVAFKGVHSLIILGAPRHQHTGKVIFTO	439
Qy	422	NTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLICAPHYGTGRCQGVSVCPPL	481
Db	440	EARHWRPKSEVRGTQIGSYFGASLCSVDVDRDGSSTDLVLICAPHYGTGRCQGVSVFPVP	499
Qy	482	RGQARWOCDAVLGEQOPGRGCAALTVLGDVNGDKLTDAVAGPGEEDNRCAVYLPH	541
Db	500	-GVRGRWCEATLHGQHPGRFQVATVLDVNGDNLADVAIGAPGEESRGAIVIFH	558
Qy	542	GTSGSGISPSHSQRIAGSKLSPRLQYFQCSLGGQDLTMDGLVBLTVGAQGHVLLRSQP	601
Db	559	CASRLIEMPSQSRVTSQSLRLQYFQCSLGGQDLTMDGLVBLTVGAQGHVLLRSPL	618
Qy	602	VLRVKAIMEPPEVARNVPECNDQVVGKXGAEVRVCLHVOKSTRDLREGQLQSVVTV	661
Db	619	LLKVELSRFAPMEVAKAVYQWERTPTVLSAGEATVCLTVHKGSDDL--GNVQGSVRY	676
Qy	662	DIALDSGPHSRVAFNETKNSRTRQTVLGLTQTCETLKLQPCNIEDPSPVILRLNFS	721
Db	677	DIALDQGLISRAIFDETKNCTLTKRKTGLGDHCEVTKLLPDCVDAVSPILRLNFS	736
Qy	722	LWGTPLSAPGNLRLVLAEDAQLFALPFFKXNCNDNICODDLSITFSMSLCLVWGG	781
Db	737	LVRDSASP--RNHLPLVAGSQDHITASLPFEKXCKQELLCEGLGISTFNSGLQVLVWGG	795
Qy	782	PREFNVTVRMDGDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEV	841
Db	796	SPELTVTVTVNNEGDSYGLVKKFYYPAGLSYRVGTQ-QPHQYPLRLACEAPAAQED	854
Qy	842	GALKSTSCSINHPIIPENSEVTNITFDVDSKASGNKLLKANVTSENMPRTNKTBPQ	901
Db	855	--LRSSSCSINHPIIFREGAKTTFMITFDVSYKAFGLDRLLLAKASSENKPDNTKTAQ	912
Qy	902	LELPKYAYVMVTSHGVSSTKYNLTASNTSR--VMHQYQVSNLQGRSLPSISLVELPV	960
Db	913	LELPKYAYVITLISQEDSTNVNFFSSSGRRQBAARVYRNVLSPKLAVRVNFWVPV	972
Qy	961	RLNQTVINDRPQVTFSENLS--TCHTKERLPSHSDPLAELKAPVYVNCISIAVCQRIQDI	1019
Db	973	LLNGAVMD--VTLSSPAQGVSCVSMKPPQNPDPFLTOIQRRSVLDGSIADCLHFRCDI	1029
Qy	1020	PFPGIQEENATLKNLSFDWIKTSHNHLAVSTABILFNDSTVFLPAGQAFVRSQTE	1079
Db	1030	PSLDIQDELDFILRGNLSFGWVSQTLQEKVLLVSEAEITFDTSVYSQPLQAEFLRAQVE	1089
Qy	1080	TKVEPEVENPLPLIVGSSVGGILLIALLITAAVLYKLGFFPKRYQXDM 1126	
Db	1090	TTLBEVTVYEPFLVAGSSVGGILLIALLITVLYKLGFFPKRYQXKEML 1136	

RESULT 7

Q9WTV4	PRELIMINARY;	PRT; 1161 AA.
AC	Q9WTV4;	
DT	01-NOV-1999 (TREMELrel. 12, Created)	
DT	01-NOV-1999 (TREMELrel. 12, Last sequence update)	
DT	01-JUN-2003 (TREMELrel. 24, Last annotation update)	
DE	Integrin alpha L.	
GN	ITGAL.	
OS	Mus musculus (Mouse).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.	
OX	NCBI_TaxID=10090;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=DBA/2J; TISSUE=Spleen;	
RA	Ma R.Z., Teuscher C.;	
RL	Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; AF065902; AAD25885.1; -.	
DR	HSSP; P20701; ILFA.	
DR	MGI; 96606; Itgal.	
DR	GO; GO:000305; C:integrin complex; IEA.	
DR	GO; GO:0004895; F:cell adhesion receptor activity; IEA.	
DR	GO; GO:0007160; P:cell-matrix adhesion; IEA.	
DR	InterPro; IPR00413; Integrin_alpha.	
DR	InterPro; IPR002035; VWP_A.	
DR	Pfam; PF01839; FG-GAP; 3.	
DR	Pfam; PF00357; Integrin_A; 1.	
DR	Pfam; PF00092; vwa; 1.	
DR	PRINTS; PR01185; INTEGRINA.	
DR	PRINTS; PR00453; VWFADOMAIN.	
DR	SMART; SM00191; Int_alpha; 5.	
DR	SMART; SM00327; vwa; 1.	
DR	PROSITE; PS00242; INTEGRIN_ALPHA; 1.	
DR	PROSITE; PS0234; VWFA; 1.	
SQ	SEQUENCE 1161 AA; 128240 MW; 865102P7B209E431 CRC64;	

Query Match	25.7%;	Score 1512.5;	DB 11;	Length 1161;
Best Local Similarity	33.7%;	Pred. No. 5.4e-103;		
Matches	394;	Conservative 219;	Mismatches 464;	Indels 91; Gaps 37;
Qy	1	FNLDTENAMTFQENA-RGFGQSVVQLGSRVVGAPQEIIVANQRGSLYQCDYSGCEP	59	
Db	24	YNLDRPTQSLAQAGRHFGYVQLIEDG-VVVGAGE---GDTGGLYHCRTSSEFCQP	79	
Qy	60	IRLOVPVEAVNMSLGLSLAATTSPPOLLACGPTVHTCSENTYVKGCLFLGSLNRQQPQ	119	
Db	80	VSLH-GSNHTSKYLGMTLATDAKGLSLADPGLSRTCDQNTYLSGLCYLFPQSLGPM	138	
Qy	120	KFPEALRGCPQSDIAFLIDSGSIIIPDPRMRELSTIMEOLKSKSLFSLMOYSEEF	179	
Db	139	QNRPAYCECMKGVYDLVFLFDGSGSLDRKDFEKLFPMDVMRKLSNTSYQFAAVQFST	198	
Qy	180	PRHPTKEP-QNNPNRSLIKPITQLLGRTHATGLRKVVRELFNITNGARKNAKILF	238	
Db	199	CRTEFTLDYVXQKNPDVLGVSQPMFLTNTPRATVYVAVHVKESGARPPATKVLV	258	
Qy	239	LLTDCGEF--GDPLGYEDVIPELDREGVRYVVLGFGDAFRSEKSRQELNTVASKPPR	296	
Db	259	IIIDGEASDKGNISAAHD-----ITRYIIGIKGHFVSQVQKTLHIPASEPVEEF	309	
Qy	297	FOANNFALKTVQNLREKIFAIEGTQSGSSSFEHMSQEGFSAAITSNGLLTVGSDY	356	
Db	310	KILDTFEKLDLFDLQRIYAEITGRQDLTSFNMELSSSGISADLSKHAVVAGVAK	369	
Qy	357	DWAGGVF-LVTSKESKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYO	414	
Db	370	DWAGGFLDLREDLQGAIFVGOEPLTSDVRGGLYGTVAWMTSRSSRPLAAGADRYQHV	429	
Qy	415	LVAMPR--QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLICAPHYGT	472	

```

Db 430 QVLLFOAPEAGGRWQKIEGTQIGSYFGCEILCSVDLDQDGEAEILLIGAPLFFGEQRG 489
QY 473 GOVSVCLPLPGQARWQCDVAVLGEQGPGRFGAALTVDGVNKGDKLTDVAIGAPCEHD 532
Db 490 GRVPTY---QRRQSLFEMVSELQDPCGPIGRFGAAITALTIDNGDLTDAVAGAPLEB- 545
QY 533 NRGAVYLFGHTSGSGSPSHQSORLAGSLSPRLQYFGQSLSGGQDITMDGLVLDLTGAQG 592
Db 546 -QAVYIFNGKPG-GLSPQSPQRIQAGVFGIRWFGSRSHGVKDLGGDLADVVVGPEG 603
QY 593 HVLLRSQPVLRKVAIMEFNPVARNVPCNDQVVGKEAG-EVRVCLHVOKSTRDLRL 651
Db 604 RVVLLSRPVDVVTLSFSPSEIIPVHEVECSYSAREEQKHGKLVKACFRKPLTPQ--F 661
QY 652 EGQIOSVVTYDLALDSGRPHSAVENETKSTRQTVGLGLTQTCETLKLQPNCTEDPV 711
Db 662 QGKLLWLSYTLQDGHWRKSLGFLPDGSHELSGNTSITP-DKSLDHFHFPICIQDLI 720
QY 712 SPIVLRLNPSLV---GTPLSAFGN-LRPLVAEDAQRLLFTALFPFKNCGNDNICQDLSI 767
Db 721 SPINVSINFLSLEEBGTFRDQKRAMQPILRPSIHTV-TKEIPFKNCGEDKCEANLTL 779
QY 768 TSPFMSLDCLVGGP-----REFNVTVVRNDEGDSYRTQVTFPPDLSTRKYSTLQN 821
Db 780 SSPARS-----GPLRLMSASLAVETWLSNGEDAYVYRLDLDPFRGLSPFKVEMLQ- 831
QY 822 QRSQGRWRLACESASSTEVSGAL-KSTSCSINEPIPPENSEVTFNITFVDVSKASLGNKL 880
Db 832 --PHSRMPVSCSEL--TEGSLTKLTKONVSPIFKAGEVSLQVWNTLNSHEDFV 887
QY 881 LKXANTSEN-NPRTNKTTFQLELPVKYAVYVMTSHGVSTKYLAFTASNTSRVMOHQ 939
Db 888 ELNGTVHCENENSLOEDNSAATHIPVLPVNLTKEQENSTLYISFTPKGPKTQGVQHV 947
QY 940 YQV---SNLQSRSLPISLVPVRLNQTVMDRPO--VTFSENLSS-----TCHTKE-RL 989
Db 948 YQVRIQPSAYDNMP-TLEALGVV-----WPHSDPPIYTVSWQTDPLVCHSELKR 1000
QY 990 PSHSDFLAELRKAPVNCYSIAVCQRIQCDIPFGIIEEFNATLKGNSLFDWIKTSHNL 1049
Db 1001 PSSE---AEQPCLPGV-----QPRCPIVF---RREILIQVGTVELSKEIKAS-STL 1045
QY 1050 LIYSTABILPNDVFTLLPQGAFAVRSOTETKVEFEVNPPLVGVSSVGLLALALIT 1109
Db 1046 SLCSLSVSFNSSKHFLHYGSKA-SQAQVLVQVDLIHEKMLHVYVLSGGVGLVLLFLIF 1104
QY 1110 AALYKLGFFPKRQYKDXM-SBGGPFGAP 1136
Db 1105 LALYKVGFFPKRLKENGHEADGGVPGSP 1132

RESULT 8
Q9R200 PRELIMINARY; PRT; 1160 AA.
AC Q9R200;
DF 01-MAY-2000 (T-EMBLrel. 13, Created)
DF 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DF 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Integrin alpha L.
GN ITGAL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spleen;
RA Ma R.Z., Teuscher C.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP065901; AA025884.1;
DR HSP; P20701; IIPA.
DR MGD; MGI:96606; Itgal.
DR GO; GO:0008305; C:integrin complex; IEA.

```

```

DR GO; GO:004895; P:cell adhesion receptor activity; IEA.
DR GO; GO:007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Intergrin_alpha.
DR InterPro; IPR002035; VWF A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWFA; 1.
DR SEQUENCE 1160 AA; 126127 MW; A33C531B139F1FAD CRC64;

Query Match 25.5%; Score 1502; DB 11; Length 1160;
Best Local Similarity 33.6%; Pred.No.3.3e-102;
Matches 392; Conservative 218; Mismatches 467; Indels 90; Gaps 36;

QY 1 FNLDTENAMTFQENA-RGFQSQVVOLOGSRVVVVGAPQBIIVAAVNOGSLVOCQDYSTGSCSP 59
Db 24 YNLDTPTQSFIAQAGRHFQVQLIEDG-VVVGAPGE---CDNTGGLYHCTSEFCOP 79
QY 60 IRLQVVEAVNMSLGLSLAATTPPOLLAGCQPTVHOTCSENTYVKGLCFGLSGNLRQQOQ 119
Db 80 VSLH-GSNHTSKYLGWTLATDAAGKSLACDPLGSLRTCDQNTYLSGLCYLPQSLGEPML 138
QY 120 KFPBALRCQPCSDPIAFILIDCGSIIIPHDERMELVSTIMEQLKSKTSLSLMOYSEE 179
Db 139 QNRPAIQCCKGKVDLVFLFDGSLDRKDFEKILFMDVNRKLSNTSYQAAVQFSTD 198
QY 180 FRIHTFKBF-QNNPNPRSLKPIITQLLGRTHATAGLRKVRELFNITNGARKNAFKILF 238
Db 199 CRTEFTFLDYVKNKNPVDLLGSLGVQPMFLLTWTFRATVVAHVFKESGARGDATKVLV 258
QY 239 LITDCEKE--GDPLAYEDVIFELDEGVIRVYLGFDGDAFRSEKSRQELNTVASKPRDHV 296
Db 259 IITDGEASDKGNISAAHD-----ITRYIIGIKGFVSVQKQKTLHIPASEPVEBEV 309
QY 297 FOANNFEALTKVQNLREKIFAIEGTQSGSSSPHEMSQEGFSAITNSGPLLSTVGSY 356
Db 310 KILDTFEKLDLFTDLQRIIVALEGTNRQDLTSFNKELSSSGISADLSKGHAVVGAGAK 369
QY 357 DWAGGVF-LYTSKHKSTFINMTRVDSMDNDVLYGAA--IILNRVQSLVLGAPRYOHIG 414
Db 370 DWAGGFLDLREDLQCATFVGQEPDTSVRRGGYLVTVAMWTSRSRPLLAAGAPRYOHVG 429
QY 415 LVAMFR--QNTGMWESNANVKTOLGAYFGASLCSVDVDSNGSTDLVLIGAPHYEOBTRG 472
Db 430 QVLLFOAPEAGGRWQKIEGTQIGSYFGCEILCSVDLDQDGEAEILLIGAPLFFGEQRG 489
QY 473 GOVSVCLPLPGQARWQCDVAVLGEQGPGRFGAALTVDGVNKGDKLTDVAIGAPCEHD 532
Db 490 GRVPTY---QRRQSLFEMVSELQDPCGPIGRFGAAITALTIDNGDLTDAVAGAPLEB- 545
QY 533 NRGAVYLFGHTSGSGSPSHQSORLAGSLSPRLQYFGQSLSGGQDITMDGLVLDLTGAQG 592
Db 546 -QAVYIFNGKPG-GLSPQSPQRIQAGVFGIRWFGSRSHGVKDLGGDLADVVVGPEG 603
QY 593 HVLLRSQPVLRKVAIMEFNPVARNVPCNDQVVGKEAG-EVRVCLHVOKSTRDLRL 651
Db 604 RVVLLSRPVDVVTLSFSPSEIIPVHEVECSYSAREEQKHGKLVKACFRKPLTPQ--F 661
QY 652 EGQIOSVVTYDLALDSGRPHSAVENETKSTRQTVGLGLTQTCETLKLQPNCTEDPV 711
Db 662 QGKLLWLSYTLQDGHWRKSLGFLPDGSHELSGNTSITP-DKSLDHFHFPICIQDLI 720
QY 712 SPIVLRLNPSLV---GTPLSAFGN-LRPLVAEDAQRLLFTALFPFKNCGNDNICQDLSI 767
Db 721 SPINVSINFLSLEEBGTFRDQKRAMQPILRPSIHTV-TKEIPFKNCGEDKCEANLTL 779
QY 768 TSPFMSLDCLVGGP-----REFNVTVVRNDEGDSYRTQVTFPPDLSTRKYSTLQN 821

```

```

Db 780 SPARS-----GPIRLMSASLAVENWTLNSGEDAYVWRLDLPFRGLSFRKVENLQ- 831
QY 822 QRSQSRWRLACASSTVSGAL-KTSCSINHPIPPENSEVFNITFDVDSKASLGNKL 880
Db 832 --PESRMPVSCBELL--TEGSSLLTKLKNVSPPIFKAGQVSLQVWMTLLNSSWEDFV 887
QY 881 LLKANTISEN-NMPTNKTFQLELPVKVAVYVWVTSBGVSTKYLANFTASENSTRVMOHQ 939
Db 888 ELNGTVHCENENSLOEDNSAATHIPVLPVNLTKEQENSTUYISFTPKGPKTQOVQHV 947
QY 940 YQVSNLQGRSLPISLVFLVRLNQTINDRPO-----VTFSENLS-----TCHKE-RLP 990
Db 948 YQV-----RIQPSAYDHNMT-LEALGVPRPHSEDLITVWSVQTDPLVTCHESEDLKRP 1001
QY 991 SHSDPLAELKAPVAVCSIAVCORIQCDDIPFGIQEBFNATLKNLSFDWYIKTSENHLL 1050
Db 1002 S-----SBAEPLPGV--OFRCPVIF--RWEILQVGTVELSKEIKAS-STUS 1045
QY 1051 IVSTABELFNDVFTLLPGGAGFVRSQTETKVBPFFVNPPLPIVSGSVGGLLALLALITA 1110
Db 1046 LCSSLSVFSNASKHFLYGSKA-SEAQLVKVLDLHEKMLHYVVLGGIGLVLFLFL 1104
QY 1111 ALYKLOFFKQYQDM-SGEGPGGAP 1136
Db 1105 ALYKVGFFKRLKXERMEADGGVNGSP 1131

RESULT 9
Q98TF1 PRELIMINARY; PRT; 1196 AA.
AC Q98TF1;
DT 01-JUN-2001 (T:EMBLrel. 17, Created)
DT 01-JUN-2001 (T:EMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (T:EMBLrel. 25, Last annotation update)
DE Cdi1-1.
GN CIAL.
OS Cyprinus carpio (Common carp).
OC Rukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Cyprinus.
OC NCBI_TaxID=7962;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Peritoneal exudate cells;
RA Kimura M., Nakao M., Miura C., Fujiki K., Yano T.;
RT "Molecular cloning of a leukocyte integrin from the common carp.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB048536; BAB39134.1; --
DR HSP; P20701; ILPA.
DR GO; GO:0008305; C.integrin complex; IEA.
DR GO; GO:0004190; P.aspartic-type endopeptidase activity; IEA.
DR GO; GO:0004895; P.cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P.cell-matrix adhesion; IEA.
DR GO; GO:0006508; P.proteolysis and peptidolysis; IEA.
DR InterPro; IPR001969; Asparticase AS.
DR InterPro; IPR004113; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS00234; VWF_A; 1.
SQ SEQUENCE 1196 AA; 132477 MW; 9369C607BDCA53B CRC64;

Query Match 23.5%; Score 1382; DB 13; Length 1196;
Best Local Similarity 31.0%; Pred. No. 2.9e-93;
Matches 369; Conservative 226; Mismatches 451; Indels 144; Gaps 39;
QY 1 FNLDTENAMTFQENARG-FCQSVVQLQ-GSR--VVGAPQEIIVAAQNRGLSQCDYSTGS 56

```

```

Db 32 FNIDTEHFLRENGAPEDFFGYSVYQTFBNRKRQIIIVGAPLE---GNSTGEMISCTADLQS 88
QY 57 CEPITELQVP----VEAVNMSLGLSLAATSPQLLACGFTVHOTCSENTYVKGCLFLPQS 112
Db 89 CQ--RLQRPGBSVRPFQMSAAVSSAALTS-----CSYPFPHCDGNSYLVNGVCYQFS 140
QY 113 NLQPOQKPFPEALRCPQEDSDIAPLIDSGSIIIPHDFRMKELVSTIMEQLKSKTLPFS 172
Db 141 SL-QAVSNFTAYQCSKREYNLVFLDGSSMKTVDFEMNKNFKIDIMKLSNKSIFPA 199
QY 173 LQVSEERIRIHTPFKEFQNNPNRSLIKPIITQLLGRTHATGLRKVVRELFN-ITNGARK 231
Db 200 AVQPSDTRVTFDFNDYQSGSAEELKME-THMSLITNTHKAIIDVILKNLLNSMLSGADS 258
QY 232 NAFKILFLLTGRKCKDPLGYED--VIPELDEGVIRVVLGFDGAFREKSRQELNLTVAS 289
Db 259 KAQKALVITD---GDFSDNDDYVNLKCKDEQNTILRYIIVG-----KVDLIEITQLAS 309
QY 290 KPRDRHVQANFALKTVQOLREKIPAIETOTGSSSSFEHEMSQRFSAAITSNGLP 349
Db 310 EFKRNTTFIKDYSGLKGLDNLQKKIYNIEGSDVAQGRDRQKELSQSGFSVYQBSVI 369
QY 350 LSTVGSYDWAGVFLYTSKEKSTPFINMTRVSDMH-DAYLGYAAAIILNRRVQSLVGLNP 408
Db 370 VGSVGSNDWRGALYEVTC-SGSDFKETIIPAVNCKDSYMGYSTVLGHRRGVSLFSGAP 428
QY 409 RYCHIGLVAMFRONTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSDLDVLIGAPHYVE 468
Db 429 RAETGLTGLTFLTKQNTWYTSNNGEIQISYFGASLLDVSQSDGSDFLVGLAPLFYQ 488
QY 469 QTRGGQSVCPLPQGRARMQCDVLYGRQ-----GQWGRPGAALTVLGDNVNGKLTDV 523
Db 489 SQ-----PRTEGRLYVYSLSEOKYFQKTLNVQSQTTRGFAASVASIKDLNGDGLSDV 540
QY 524 ATGAPGBEDNRGAVLPHGTSGSLSPSHS-ORLAGSKLSPRLQVFGOSLGGQDLTWG 582
Db 541 AVGAPLE--NEGVTIYIGDRTHGHNELTFORISVQSVLPGLQFQVSLTGQDMNDN 598
QY 583 LVDLTVGAQGVHLLRSQVLRKAIMFNPREVARNVPECNDQVVKGEAGEVAVCLHV 642
Db 599 LTVDIVGAGGIVLLKARPMSVSAQLSPSKISLNYFECFGS--NAPNAPLTSCTVP 656
QY 643 QKSTEDRLREGQISV--VTVDLALDSGRPSRAVFNETKSTR--RQTQVLGLTQTCET 698
Db 657 TERTSS--TGSLEKLLNVSLNLDVVDVVGMSRGFPDQSSVSSRTLQSVLLDSSGSCFN 713
QY 699 LKQLPNCIBDPVSPVILRLNPLSLVGTPLSAFGLNRPVLAEDAQRLFTALPFPKNCND 758
Db 714 PSIPLMLRCVADTVSPLKIRMNFS--QTEMLS-GNSVAVLDVHSRTEENVVFPQNCNSN 770
QY 759 NTCQDLSITFSFMSLDCLVGGPREFNWTVTRVDGDSRTQTQVTFPDLDSYKYST 818
Db 771 NSCVADLKLNFSTN-NTLVVENQAHTVQVSLANPGGDSYNTSVTLHYPGISLSPDA 829
QY 819 LQNRQSRSWRLACASSTVSGALKSTCSINHPIPPENSEVFNITFDV---DSKAS 875
Db 830 IKPSR-----SSCGDRDSGATNFTSIDLPVRSCTTQFLGTFRVMKWDNDS 881
QY 876 LGNKLKLLKANVTSNNMPTNKTEPQLELPVKVAVYVWVTSBGV-STKYLAPTASENSTR 934
Db 882 NRWEIMITANSNNMNM---SDTEVRRESVPVQFAVDLAIISLVAEDSVTYMNFSLDROPK 938
QY 935 VQCHOYQVSNLQGRSLPISLVFLVPR-----LN 963
Db 939 PLNITYKVENSGRDLFPVSVTLTLPQCPHVLTPHTFSMEVHVSPISSVHQIIMCLLN 998
QY 964 QTVIWDREPQVTFSENLSSTCHTKERLPKSHSDPLAELKAPVAVVNCISVACQRIQ-----C 1017
Db 999 XHLLFFSPSLSAVQQRSTGWSLREVEK-SQF--DLNKSAAVHFNLTADARLQNVKYES 1055
QY 1018 DIPFGIOEB--FNATLKNLSFDWYIKTSENHLLIVSTAILFNDSVFTLLPGQAFVR 1075

```

```
Db 1056 KYSFPEFRKDNVPSISAEIYNNTSLYNQTS-----SELKYNPH-----R 1094
QY 1076 SOTETKVEPPEVNPPLIV-GSSVGGILLALLITAAALYKLGFFKQYKD 1124
Db 1095 SOTETKVE-FVVPSPSLMLIVCTCAVGGFFFLIILFLLKCGFFKRNRPD 1143

RESULT 10
Q96HB1 PRELIMINARY; PRT; 1086 AA.
AC Q96HB1;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue: lymph;
RA Scarsberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC008777; AA08777.1; -.
DR GO; GO:000305; P:cell adhesion complex; IEA.
DR GO; GO:004895; P:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR SMART; SM00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; vwa; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWF_A; 1.
KW Hypothetical protein.
SQ
SEQUENCE 1086 AA; 119223 MW; P6FF2546E8C632F9 CRC64;
Query Match 22.8%; Score 1341.5; DB 4; Length 1086;
Best Local Similarity 32.0%; Pred. NO. 2.5e-90;
Matches 374; Conservative 186; Mismatches 434; Indels 175; Gaps 37;
QY 1 FNLDITENAMTFQ--ENARFGOSVVOLOGSVVVGAPQIBVAANORGSLYQCDYSTGSC 58
Db 26 YNLDVGRGSPSPPRAGRHFYRVLVQ-CNGVIVGARGE--GNSGSLYQCSGTGHL 81
QY 59 PIRLQVPEAVNMSIGLSLAATSPOLLACGTVHQTCSNTYVKGCLFPGSNLRQP 118
Db 82 PVTER-GSNYTSKYLGMTLA--TDP----- 103
QY 119 QKPEALRGCPQSDSIAFLIDGSGSIIPDFRMKELVSTIMEQLKSKTFLSLMOYSE 178
Db 104 -----TDS-----ILFAAVQEST 117
QY 179 BFRHFTTFKFPNNPRLSIPITOLLGRTHATGLRVKRVRELNITNGARKNAFKILF 238
Db 118 SYKTEFDPSDYVKKRQDPDALLKHVHMLLTNTFGAIVYVATEVFEELGASPDATKVL 177
QY 239 LITDGEKGDPLGYEDVIVPELDRGVIRVYVVGFGDAFRSEKSRQELNTVASKPRDHVQ 298
Db 178 IITDGE--ATDGNDAAND-----IIRYIIGIKHFQTKESCETLHKFASKPASEFVK 230
QY 299 ANNFEALKTQVQLREKIFAIGTQTGSSSPEHEMSQGFSAATISNGPLLSITVGSYDW 358
Db 231 LDTFEKLDLFTLQKKIVIEGTSKQDLTSFNMELSSSGISADLSRGHAVVAVGAXDW 290
QY 359 AGGVF-LYTSKSKSTINNRVDSNDNDAYGVAAA-IILNRVQSLVGLGAPYOHIGLY 416
Db 291 AGGFLDKADLQDDTFIGNELPTPEVRAGLYGVYVTLWLPSPROKTSLLASGAPRYOHMGRV 350
```

RESULT 11

```
Q98TF0 PRELIMINARY; PRT; 1187 AA.
AC Q98TF0;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CD11-2.
OS Cyprinus carpio (Common carp).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Cyprinus.
OX NCBI_TaxID=7962;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue: Peritoneal exudate cells;
RA Kimura M., Fujiki K., Nakao M.;
RT "Molecular cloning of a leukocyte integrin from the common carp."
```

RL	Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; AB048537; BAB39135.1; --	
DR	HSP; P20701; ILPA	
DR	GO; GO:0008305; C:integrin complex; IEA.	
DR	GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.	
DR	GO; GO:0004895; F:cell adhesion receptor activity; IEA.	
DR	GO; GO:0007160; P:cell-matrix adhesion; IEA.	
DR	GO; GO:0006508; P:proteolysis and peptidolysis; IEA.	
DR	InterPro; IPR001969; Aspartate AS.	
DR	InterPro; IPR000413; Integrin_alpha.	
DR	InterPro; IPR002035; VWF_A.	
DR	Pfam; PF01839; FG-GAP; 3.	
DR	Pfam; PF00357; Integrin_A; 1.	
DR	Pfam; PF00092; vwa; 1.	
DR	PRINTS; PRO1185; INTEGRINA.	
DR	PRINTS; PRO0453; VWFADOMAIN.	
DR	SMART; SM00191; Int_alpha; 5.	
DR	SMART; SM00327; VWA; 1.	
DR	PROSITE; PS00141; ASP_PROTEASE; 1.	
DR	PROSITE; PS0234; VWF; 1.	
SQ	SEQUENCE 1187 AA; 131778 MW; 85EDC7CA8B6B1C59 CRC64;	
	Query Match 22.6%; Score 1329.5; DB 13; Length 1187;	
	Best Local Similarity 31.0%; Pred. No. 2.3e-89;	
	Matches 369; Conservative 210; Mismatches 456; Indels 157; Gaps 43;	
QY	1 FNLDTENAMTFQENARG-FQGSVVQLQ-GSR--VVTGAPQEIIVANORGLYQCDYTGSG 56	
DB	32 FNIDTEHLEFNGTPEDFGYSYVQTEFGNKRKIIVGAPLE---GNSAGEMYCTADLOS 88	
QY	57 CEPIRLQVP---VEAVNMISLGLSLAATSPQQLACGPTVHQCSENVTKGLCFLFGS 112	
DB	89 CK--RLORPQSESVRFPGMSAAVSSAALTS-----CSPYFAHECDGNSVLNGVCYQFNS 140	
QY	113 NLRQOPKQPEALRGCGQEDSDIAFLIDGSGSIIPHDFRMKELVSTIMEQLKSKTLFS 172	
DB	141 SL-QAVSNFTAAVQECSEKREVLVFLPDGSSSKAVEFDMNKNFIKDVNMKLSNNSIKEA 199	
QY	173 LMOYSEFRHFTKFFQNNPERSLKPTQLLGRTHATGLRKVVRIEN-ITWGARK 231	
DB	200 AVQFTEIRTVFNDYFQNGSAEKLKCE-RHMKSLFTYKAINYLVKNVLSVSSGADP 258	
QY	232 NAFKILFLLDGKFGDPLGYED--VPELDRGVIRYVLGPDGAFRSEKSRQELNVTAS 289	
DB	259 NAKKALVIITD---GDPSNDVYINLINCQENILRYIIGV-----KYDLTTLTQLAA 309	
QY	290 KPRPDHVFQANNPEALKTVQNLREKIPAFECTQTGSSSFHEMSQEGESAITSNGPL 349	
DB	310 EPKLNNTFYTOEYNGLGLDLNIOKKIYNTGSKAHRDRQKELSGSGSVVYQESVI 369	
QY	350 LSTVGSYDAGGVFLYT---SKEKSTFINNTRYDSDMDAYLGAAAILNRVQSLVLG 406	
DB	370 VGSVSGNDKAGLYEVWVGSGSKFRQTEITDPVAVN---KDSYMGYSTVLGRHGVSLFSG 426	
QY	407 APYQHILGLVAMFRONTGMESNANVKGTOIGAYFGASLCSVDVDNGSTDLVLIGAPHY 466	
DB	427 APRAEHTGLVLTFRNESTVTVMRNINGEQIGSYFGASLLDSDGSDPFLVAGAPLF 486	
QY	467 YE-QTRG-GQVSVCPFPRGQARWQCDVAVLGSQ--GQPMGRGAALTVLGDVNGDKLTD 522	
DB	487 YQSQPRAEGLYVTL-----SEYQSKTQSTTGRFATSLASLXDLNGGLSD 535	
QY	523 VATGAPGEENRGAVLYFHTSGSGSPSHS-ORLAGSKLSPRLQFGQSLSGQDLTWD 581	
DB	536 VAVGAPLE--NEGVTYTLGDGTHGHNPEHAPQIPARSVLPGLQQFGVSLSGQMDMND 593	
QY	582 GLVDLTVGAQGHVLLRSQPLAVKAIMENPREVARNPECDQVQVKGKEAVEVCLH 641	
DB	594 NLPDIVIGTGGIVLLNARPMVMSVQLSFNPHEISLNYFECPS--NAPNAELTSCFT 651	
QY	642 VQKSTRDLREGQIQSV--VYTDALDSDGRPHRAVFNETKNSTR--RQTVLGLTQTC 697	
DB	652 VTERTS--TGSLEKLVNLSMLNLDVVRGMRGPFDPMDSSRTLQQSVLLDSGSGS 708	

QY	698 TLKLOLPNCIEDPVSPVILRLNFS---LVGTPLSAFGLNLRPVLAEADAQRLFTALPPPEK 753	
DB	709 NESIFMLRCVADTVSPFKIRMFSTQMLSGNSLAVL-DIQSRTEYVEVL-----PQR 761	
QY	754 NCGNDNICODDLISITFTSMSLDCLVVGGRPREPNVTVTVRNDEGDSYRTQVTFPPPLDLSY 813	
DB	762 NC-NGNSCVADLKLNFSTN-DTLVVENQAHFTVLVSLANPGDDSYNTSIVLHYPEGLSL 819	
QY	814 RGVSTLQNRQSRQSWRLACEGASSSEVSGALKSTSCSINHPIFPENSEVTFNITFDV--- 870	
DB	820 SKFDAIKFSRTK-----SSCGDRDSGATNRTTCSINLPVYRSGTTQFGLGTRVTKW 871	
QY	871 DSKASLGNKLLKANVTSENMMPTNKTEFOLELPKYAVVMVVTSHGV-STKYLNFTAS 929	
DB	872 DYDWSDRMEMTITANSNNNGM---SDMSVRRSIPVQFAVELAISLVAEDSVTYLNFSL 928	
QY	930 ENTSRVMQHOYQVSNLQORSIPISLFLVPLVPLNQTIVDRPQVTFSENLSSTCHTKERL 989	
DB	929 DRGPRPLNIYKVNLGKLPVSVTLSPQC-----TTHVTILTPhnFSMQ 974	
QY	990 PSHSDFLAELRKAPVVC-----SIAVQORIQCDDIPFGIOEE 1027	
DB	975 EVHHSFISSYHQ--IIMCLLNKHLFPSPELSAVQTRTGRSLWVC-----VSSISTGEI 1026	
QY	1028 FNATLKGNL-----SFDWYIKTSHNHLIVSTAEILFNDSVFTLLPQ-QGA 1072	
DB	1027 FRSSV--NLMAEAVLQNVKEYSKSYF-EFRDRHVFNIS-AELNFTSRYNQSTGLKTN 1082	
QY	1073 FVRSQTEYKVPFEPVNPPLPLTVGSGVGLLLALITAALYKLGFFKQYKD 1124	
DB	1083 PHRSQTEYKVPFVPIPSRMLIVCTGAVGSGFFLLIILLKCGFFKMRPD 1134	

RESULT 12	
Q8HZVO	PRELIMINARY; PRT; 927 AA.
AC	Q8HZVO;
DT	01-WAR-2003 (Tremblrel. 23, Created)
DT	01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT	01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE	Lymphocyte function-associated antigen 1 (fragment).
OC	Bos taurus (Bovine).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC	Suidae; Bovinae; Bos.
OX	NCBI_TaxID=9913;
RN	[1]_TaxID=9913;
RP	SEQUENCE FROM N.A.
RA	Thumbikat P., Kannan M.S., Maheswaran S.K.;
RT	"Sequence of the alpha subunit of bovine lymphocyte function-
RT	associated antigen 1."
RL	Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AF440778; AAN63636.1;
DR	PIR; A32039; A32039.
DR	GO; GO:0008305; C:integrin complex; IEA.
DR	GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR	GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR	InterPro; IPR000413; Integrin_alpha.
DR	InterPro; IPR002035; VWF_A.
DR	Pfam; PF01839; FG-GAP; 3.
DR	Pfam; PF00092; vwa; 1.
DR	PRINTS; PRO1185; INTEGRINA.
DR	SMART; SM00191; Int_alpha; 4.
DR	PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR	PROSITE; PS0234; VWF; 1.
FT	NON_TER 1
FT	NON_TER 927
SQ	SEQUENCE 927 AA; 102523 MW; 02E2CF09917081EC CRC64;

Query Match 21.5%; Score 1264; DB 6; Length 927;	
Best Local Similarity 34.7%; Pred. No. 1.1e-84;	
Matches 339; Conservative 165; Mismatches 388; Indels 86; Gaps 28;	


```
QY 722 LVGTPLSAFQNLPAVLAEDAQRLFTALP--PFFKNCNDNICODDLISITFSFMSLCLVY 779
Db 782 F-QTSEERRNHPNFI--DHYKESALFQLPYKCKKVKFCIABITQITTAISQOD-LVY 837
QY 780 GGPREFNVTVRNDGEDSTRTQVTPFPDLDSYRKVSTIQNORSORSWPLACESASTE 839
Db 838 GITKEVTMNISLNGSDSYMTNANLNPRLQFKKI-----QKPLSPDIQDDPKPV- 890
QY 840 VSGALKSTCSINHPIPPENSEVTNFTPDVDSKASLGNKLLKANVTSENMPRTNKE 899
Db 891 --ASVLVNMCKIGHPII-KRSSVNVSVTWQLESIIPNRTADITVTSNKSLSARETH 947
QY 900 FOELPVPKAYVMVTVSHGVSTKYLNFTASENTSRVMQHOYQVSNLQORSPLISLVLVP 959
Db 948 --SLOFRHAFIAVLSR--PSVMYMN--TSQSSSDHKKEFFNFVHGENHFGAVPOLQICVP 1000
QY 960 VRLNQTVWRDPQVTSSENLSST-CHTKERLPSHSDFLAELRKAPVNVNSIACQRIQCD 1018
Db 1001 ITRDL-----QIIRVGHUKTQAHY-----ECTQSQPTCGSDPVQNV--- 1039
QY 1019 IPFFGIQEFNATL-----KGNLSFDWYIKTSHNHLII--VSTABIL-----FNDSVFTL 1066
Db 1040 -----EWHSVICAITSKENVVAAEISMGHTKQLLRDISLQILGHBISFNKSLYEG 1092
QY 1067 LPOGAFVRSQTKVEFP-----EVNPLPLIVGSSVGGHLLALITALYKLGPPKROY 1122
Db 1093 ENAE-----NHRKTIVFLKBEKPHSLPLITIGSSIGGLLVVIAILFRCGFPKRY 1146
QY 1123 KDWMS 1128
Db 1147 KQNLN 1152

RESULT 15
Q7TOC3 PRELIMINARY; PRT; 1188 AA.
AC Q7TOC3
DT 01-OCT-2003 (T-EMBLrel. 25, Created)
DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE All integrin.
GN ITGAI1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Tiger C.-P., Popova S.N., Velling T., Ekman S., Forsberg E.,
RA Gullberg D.;
RT "allb1 integrin is important for mesenchymal cell function:
RT elimination of allb1 leads to dwarfism."
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Johansson M., Popova S.N.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY124460; AAM62130.1; --
KW Integrin.
SQ SEQUENCE 1188 AA; 133067 MW; 6982BF20D75E4271 CRC64;

Query Match 17.7%; Score 1038.5; DB 11; Length 1188;
Best local similarity 28.1%; Pred. No. 1e-67;
Matches 349; Conservative 182; Mismatches 496; Indels 213; Gaps 46;

QY 1 ENLDTENMTFO-ENARGCQSVVQ--LOGSR-VVVGAPQEIIVANQORSLVCCDYSTGS 56
Db 23 FNMDTNRPRVIAGPSAFTGTVQOHDISGKWLWVGAPMETNMGHGTGDKVYKCPVTQN 82
QY 57 CEPRL-----QVPVEAVMSIGLSLAATSPQQLACGPTVHQTCSNTYVKGICLFL 110
Db 83 CTKNLGRVTLNVSERKDNMRIGLSLATNPXDNSPLACPLWSHCGSSYYTTGMSRV 142
```

```
QY 111 GSNLRQOPKQPPALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMBOLKK--- 166
Db 143 NENFRESKTVAP-ALQRC-QTMDIVIVLDGNSIYP-----WVEVQHLINILKKFYIG 195
QY 167 -SKTSLPMQVSEBFRHFTPKBFONNPNRSLIKPIITOLLG-RTTATAGLARKVRELFN 224
Db 196 PQGIQVIGVQYGEDAVHEFLNDYRSVKDVVAAASHIEQRGGTETRTAFGIPARSEAFQ 255
QY 225 ITNGARKNAFKILFLLTDGKEKFGDPLGYEDVPELDRGVIY---VLFGD--AFRSEK 279
Db 256 --KGRKGKAKVNIWITDGSHPD-DEKVIROSEKKNVTYAVAVLGYNRRGINPET 312
QY 280 SQEELNTVASKPRPHVFOANNFEALKTVQNLREKI FAIEGTOTGSSSPHEMSQBP 339
Db 313 FLNETKIYASDDDKHFFNVYDEAALKDIDVALGDRIFSLEGTNK-NETSFGLEMSQTF 371
QY 340 SAAITSNGELSTVGSYDWAGVFLYTSKEK-----STFINNTRVDSMDNDAVLGAAAI 394
Db 372 SHVVVEDGILLGAVGAYDMNGAVLKEBTSAGRVIPHRESYLKEPPELKNHVAIYLGTVTS 431
QY 395 ILNRV-OSLVLAGAPRYQIHGLVAMP-RQNTGMWESNANVKGTQIGAYFASLCSYDVS 452
Db 432 VVSSRQRYVAGAPRFNHTGKVLFSMENNRSITIHOALRGEQIGSYFGSITSVDND 491
QY 453 NGSTDLVLIGAPHYEQTR-GQNSVCELPGRQARWQCDVLYGEOGQWGRFGAALTV 511
Db 492 DRVTVLLVGAEMYSEGREKGVVYNL---RQNRVYNGTLKDSHSHYQNFARFGSCIAS 548
QY 512 LGDVNGDKLTVAI GAPGEEDNRGAVYLFHGTSGSIGSPSHSQRIAGSKSLPRLOVFGOS 571
Db 549 VQDLNQDSVNDVVGAPLEDSEHGAIYIFHGQFQ-TNILLKPMQRTASELAPLOHFGCS 607
QY 572 LSGGQDLTWGLDVLTVGAQGHVLLRQPVLRVKAIMEFNPREVARNVF--ECNDQVVK 629
Db 608 IHGQLDLNEDGLVDLAVGALGNVAVLWAPVQVINASLHFPEPSKI--NIFHDC----- 659
QY 630 GKEAGEVRVCL-----HVOKSTRDLREGQIQSVTVYDLDALDSGRPHRAVF 676
Db 660 -KRNCRDATCLAAFLCFPIPIFLAPHQATVG-----IRNATWDERRYVRAHL 708
QY 677 NETKNS-TRRQTVLGLTQCTETLKLQPLNCIEDPVSPVLRNLNPSLVGTPLSAGFNLRP 735
Db 709 DEGGQFTNRVALLSSGQEHQCRINFHVL-DADYVVKPVAFSVEYSLEDP-----DNRP 761
QY 736 VLAEDAQRLLFTALPPEKNCNDNICQDDL-----HVOKSTRDLREGQIQSVTVYDLDALDSGRPHRAVF 768
Db 762 MLDNGWPTTLRVSVFVWNGCNEDEHCVPELVLDARSDLPTAMEYCORVGRPAQCCSTT 821
QY 769 PSFMSLDCLVGGPREFNVTVVRNDGEDSYRTQVTPFPDLDSYRKVSTIQNORSQSW 828
Db 822 LSFDTTVFIESTRRVAVAEATLENRGENAYSAVLNISQSENILQF--ASLIQKDDSDNS- 878
QY 829 RLACESASSTVSGALKSTCSINHPIPFENSEVTNFTPDVDSKASLGN-KLLIKANVT 887
Db 879 -IECVNEER-----RLHKVCNVSPYFFRAKAVAPRLDPFESKSVFLHLLQIHGAGSD 932
QY 888 SENNMPRTNKEFQLELPVKVAVVMVTVSHGVSTKYLNFTASENTS-----RVWQHQ 939
Db 933 SHEQDSTADDNTALLRFLHYEADVIFTR---SSLSLHFEVKANSLSYDYGIPFPCV 989
QY 940 YOVSNHIGORSPLISLVFL---VPV-----RL-----NOTVIMWRP 971
Db 990 FKVQNLG--PFIHGVWVKKITVPIATRGNNRLLMLRDFTDQNTSCNIGNSTYRSTP 1047
QY 972 QVTFSENLSSTCHTKERLPSHSD---FLAELRKAPVNVNSIACVQRIQCDIPFGIQSEF 1028
Db 1048 T---EEDLS---HAPQNHNSDVVSIIICNRLAP-----SSET 1080
QY 1029 NATLKNLSFDWYIKTSHNHLIIYSTABI---LFDNSVFTLLPQCGAFVRSQTKVEFP 1085
Db 1081 SPYLVGNLWLTSLKALKYRSLSKITVNAALQRFHSPFI-----REEDPSRQVTP 1130
QY 1086 EVPN-----PLPLIVGSSVGGHLLALITALYKLGFPFK 1119
```

Db 1131 EISKQEDWQVPIWIIIVGSLGGLLLLALZVLALWKLGFK 1170

Search completed: June 7, 2004, 17:16:16
Job time : 46.4889 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 16:54:39 ; Search time 49.4608 Seconds
(without alignments)
6495.175 Million cell updates/sec

Title: US-09-902-481B-3

Perfect score: 5879

Sequence: 1 FNLDTENAMTEQENARGFQ.....FKROYKDMSEGGPPGAEPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A.Geneseq 29Jan04:.*
1: Geneseqp1980s:.*
2: Geneseqp1990s:.*
3: Geneseqp2000s:.*
4: Geneseqp2001s:.*
5: Geneseqp2002s:.*
6: Geneseqp2003as:.*
7: Geneseqp2003bs:.*
8: Geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5839	99.3	1153	2 AAW65090	Human Bet
2	5839	99.3	1153	2 AAB07360	Human CD1
3	5839	99.3	1153	5 AAU80252	Human int
4	5839	99.3	1153	5 ABG61469	Human Bet
5	5839	99.3	1153	5 AAO14428	Integrin
6	5839	99.3	1153	7 ADD25615	Binding d
7	5829	99.1	1153	2 AAR04136	Alpha sub
8	3437	58.5	1163	2 AAR07120	p150.95 a
9	3423	58.2	1163	3 AAW65091	Human Bet
10	3423	58.2	1163	3 AAB07361	Human CD1
11	3423	58.2	1163	5 ABG61470	Human Bet
12	3421	58.2	1163	6 ABOU7406	Protein d
13	3388	57.6	1161	2 AAR78166	Human bet
14	3388	57.6	1161	2 AAW23049	Human bet
15	3388	57.6	1161	2 AAW57491	Human bet
16	3388	57.6	1161	2 AAW65089	Human Bet
17	3388	57.6	1161	2 AAW72825	Human alp
18	3388	57.6	1161	2 AAW73342	Human alp
19	3388	57.6	1161	2 AAB07359	Human alp
20	3388	57.6	1161	5 ABG61468	Human Bet
21	3372.5	57.4	1161	2 AAW23064	Human Bet
22	3372.5	57.4	1161	2 AAW65106	Human Bet
23	3372.5	57.4	1161	2 AAW72837	Human alp
24	3372.5	57.4	1161	2 AAW73343	Human alp
25	3372.5	57.4	1161	3 AAB07376	Human alp

26	3372.5	57.4	1161	5 ABG61485	Human Bet
27	3215.5	54.7	1161	2 AAR78169	Rat alpha
28	3213.5	54.7	1161	2 AAW23062	Rat beta
29	3213.5	54.7	1161	2 AAW60004	Rat alpha
30	3213.5	54.7	1161	2 AAW72824	Rat alpha
31	3213.5	54.7	1161	3 AAB07374	Rat alpha
32	3213.5	54.7	1161	5 ABG61483	Rat Beta2
33	3206.5	54.5	1161	2 AAW65104	Rat beta-
34	3206.5	54.5	1161	2 AAW73345	Rat alpha
35	3201	54.4	1161	2 AAW23061	Mouse bet
36	3201	54.4	1161	2 AAW60003	Mouse alp
37	3201	54.4	1161	2 AAW65103	Mouse bet
38	3201	54.4	1161	2 AAW72836	Mouse alp
39	3201	54.4	1161	2 AAW73347	Mouse alp
40	3201	54.4	1161	3 AAB07373	Mouse alp
41	3201	54.4	1161	5 ABG61482	Mouse Bet
42	3197	54.4	1161	2 AAR78168	Mouse alp
43	3192.5	54.3	1151	2 AAW23059	Rat beta
44	3192.5	54.3	1151	2 AAW60001	Rat alpha
45	3192.5	54.3	1151	2 AAW65101	Rat Beta-

ALIGNMENTS

RESULT 1
AAW65090
ID AAW65090 standard; protein; 1153 AA.

XX AC AAW65090;

XX XX 28-SEP-1998 (first entry)

XX DE Human Beta-integrin CD11b subunit protein.

XX KW Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;
XX KW type-I diabetes; atherosclerosis; multiple sclerosis; asthma;
XX KW lung inflammation; acute respiratory distress syndrome; CD11b subunit;
XX KW rheumatoid arthritis.

XX OS Homo sapiens.

XX XX US5728533-A.

XX XX 17-MAR-1998.

XX PF 07-JUN-1995; 95US-00485618.

XX PR 23-DEC-1993; 93US-00173497.

XX PR 05-AUG-1994; 94US-00286989.

XX PR 21-DEC-1994; 94US-00362652.

XX PA (ICOS-) ICOS CORP.

XX PI Van Der Vieren M, Gallatin WM;

XX DR WPI; 1998-206565/18.

XX PT Screening assay for modulators of integrin binding - using immobilised or
XX PT labelled alpha-d polypeptide, useful for, e.g. treating type-I diabetes.

XX XX Example 5; Fig 1A-D; 106pp; English.

XX This sequence represents a human beta-integrin CD11b subunit which is
XX used to describe a method for identifying compounds that modulate the
XX interaction of the beta-integrin alpha-d subunit with a binding partner
XX of alpha-d which involves contacting an alpha-d polypeptide with an alpha
XX -d binding partner, one of which is immobilised and the other of which is
XX labelled, in the presence of a test compound, and determining if the
XX compound affects binding between the alpha-d polypeptide and alpha-d
XX binding partner, where the alpha-d polypeptide is alpha-d or its fragment
XX comprising the cytoplasmic, transmembrane or extracellular domain of
XX alpha-d. Compounds that modulate alpha-d binding could be used to treat

CC diseases such as type-I diabetes, atherosclerosis, multiple sclerosis,
CC asthma, psoriasis, lung inflammation, acute respiratory distress syndrome
CC and rheumatoid arthritis
XX
SQ Sequence 1153 AA;

```
Query Match          99.3%; Score 5839; DB 2; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGVSQVVLQGSRRVWVGAPOEIVAAANORGSLYQCDYSTGSCPEI 60
Db      |||
QY 17 FNLDTENAMTFQENARGFGVSQVVLQGSRRVWVGAPOEIVAAANORGSLYQCDYSTGSCPEI 76
Db      |||

QY 61 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHQCSTENYVYKGLCFPGSNLRQPOK 120
Db      |||
QY 77 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHQCSTENYVYKGLCFPGSNLRQPOK 136
Db      |||

QY 121 FPEALRGCPQSDSDIAFLIDGSGIIPHPFRMKELVSTIMEOLKSKTLFSLMOYSSEF 180
Db      |||
QY 137 FPEALRGCPQSDSDIAFLIDGSGIIPHPFRMKELVSTIMEOLKSKTLFSLMOYSSEF 196
Db      |||

QY 181 RIHFTKEFQNNPRSLKPIITQLLGRTHATGLRKVVRELFNITNGARKNAFKILFL 240
Db      |||
QY 197 RIHFTKEFQNNPRSLKPIITQLLGRTHATGLRKVVRELFNITNGARKNAFKILV 256
Db      |||

QY 241 TDCEKGDPLGYEDVPELDREGVIRVVLGFGDAFRSEKSRQELNTVASKPRDHVQAN 300
Db      |||
QY 257 TDCEKGDPLGYEDVPELDREGVIRVVLGFGDAFRSEKSRQELNTVASKPRDHVQAN 316
Db      |||

QY 301 NFELAKTVQNLREKIFAIBGTQTGSSSPFHEMSQEGFSAATSNGLLSTVGSYDNAG 360
Db      |||
QY 317 NFELAKTVQNLREKIFAIBGTQTGSSSPFHEMSQEGFSAATSNGLLSTVGSYDNAG 376
Db      |||

QY 361 GVFLYTSKEKSTINMTRVDSNDAYLGYAAIILNRRVQSLVIGAPRYOHIGLVAMFR 420
Db      |||
QY 377 GVFLYTSKEKSTINMTRVDSNDAYLGYAAIILNRRVQSLVIGAPRYOHIGLVAMFR 436
Db      |||

QY 421 QNTGMESNANVAGTQIGAFYFASLCSVDVDSNGSDLVLIAGPHYEYQTRGGQVSCPL 480
Db      |||
QY 437 QNTGMESNANVAGTQIGAFYFASLCSVDVDSNGSDLVLIAGPHYEYQTRGGQVSCPL 496
Db      |||

QY 481 PRQQRARWQCDVLYEGQGFQWFRFGAALTVDVNGDKLTDVAIGAPGEENRGAVYLP 540
Db      |||
QY 497 PRQQRARWQCDVLYEGQGFQWFRFGAALTVDVNGDKLTDVAIGAPGEENRGAVYLP 556
Db      |||

QY 541 HGTSGGSISSHQRAGSKLSPLOYFGQSLGGQDLTMDGLVDTVGAQGHVLLRQ 600
Db      |||
QY 557 HGTSGGSISSHQRAGSKLSPLOYFGQSLGGQDLTMDGLVDTVGAQGHVLLRQ 616
Db      |||

QY 601 PVLVRVKAIMEFNPREVARNVFCNDQVYKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 660
Db      |||
QY 617 PVLVRVKAIMEFNPREVARNVFCNDQVYKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 676
Db      |||

QY 661 YDLALDSGRPHSAVNETKNSRROTQVGLTCTETLKLQPCNIEDPVPSPVILRLNF 720
Db      |||
QY 677 YDLALDSGRPHSAVNETKNSRROTQVGLTCTETLKLQPCNIEDPVPSPVILRLNF 736
Db      |||

QY 721 SLVGTPLSAFNLPRVLAEDAQLFTALPPFEXKNGNDNICQDDLSITFSPMSLCLVWG 780
Db      |||
QY 737 SLVGTPLSAFNLPRVLAEDAQLFTALPPFEXKNGNDNICQDDLSITFSPMSLCLVWG 796
Db      |||

QY 781 GPREFNVTVVRNDGDSYRTQVTEPFPLDLSYRKVSTLQNSORSWLACESASTEV 840
Db      |||
QY 797 GPREFNVTVVRNDGDSYRTQVTEPFPLDLSYRKVSTLQNSORSWLACESASTEV 856
Db      |||

QY 841 SGALKSTCSINPIPPENSEVTFNITFDVDSKASLGKLLKANVTSENNPRTNKTEF 900
Db      |||
QY 857 SGALKSTCSINPIPPENSEVTFNITFDVDSKASLGKLLKANVTSENNPRTNKTEF 916
Db      |||

QY 901 QLELPVKYAVYMWVTSKGVSTKYLNTASNTSENVMOHGYOYNSLQORSLSLPLVLPV 960
Db      |||
QY 917 QLELPVKYAVYMWVTSKGVSTKYLNTASNTSENVMOHGYOYNSLQORSLSLPLVLPV 976
Db      |||
```

```
QY 961 RLNQTVIMDRPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVVCNCSIAVCQRIQCIP 1020
Db      |||
QY 977 RLNQTVIMDRPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVVCNCSIAVCQRIQCIP 1036
Db      |||

QY 1021 FPGIOEEFNALKGNLSFDWYIKTSHNELLIVSTAEIILFNDSVFTLLPQOGAFVRSQET 1080
Db      |||
QY 1037 FPGIOEEFNALKGNLSFDWYIKTSHNELLIVSTAEIILFNDSVFTLLPQOGAFVRSQET 1096
Db      |||

QY 1081 KVEPPEVENPFLPIVGVSSVGGLLALITAAALYKLGFFKQYKDMMSGGPPGABPQ 1137
Db      |||
QY 1097 KVEPPEVENPFLPIVGVSSVGGLLALITAAALYKLGFFKQYKDMMSGGPPGABPQ 1153
Db      |||
```

RESULT 2
AAB07360

ID AAB07360 standard; protein; 1153 AA.

XX AC AAB07360;

XX DT 17-JAN-2001 (first entry)

XX DE Human CD11b protein sequence.

XX KW Human; macrophage infiltration inhibition; alpha d integrin;

XX KW leukocyte integrin; Leu-CAM; leukointegrin; immune response;

XX KW inflammation; leukocyte adhesion deficiency; LAD; Type I diabetes;

XX KW atherosclerosis; multiple sclerosis; asthma; psoriasis; beta2 integrin;

XX KW lung inflammation; acute respiratory distress syndrome; Crohn's disease;

XX KW rheumatoid arthritis; central nervous system injury; CD11b.

XX OS Homo sapiens.

XX PN WO200029446-A1.

XX PD 25-MAY-2000.

XX PF 16-NOV-1999; 99WO-US027139.

XX PR 16-NOV-1998; 98US-00193043.

XX PR 08-JUL-1999; 99US-00350259.

XX PA (ICOS-) ICOS CORP.

XX PI Gallatin MW, Van Der Vieren M;

XX DR WPI; 2000-387751/33.

XX Use of novel anti-alpha integrin d monoclonal antibodies to inhibit

XX macrophage infiltration and reduce inflammation at central nervous system

XX injury sites.

XX PS Example 5; Fig 1; 270pp; English.

XX Integrins are a class of membrane-associated molecules that participate

XX in cellular adhesion. Integrins are made up of an alpha subunit and a

XX beta subunit. One class of human integrins are restricted to expression

XX in white blood cells and have a common beta2 subunit: the leukocyte

XX integrins, Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins

XX have an important role in immune and inflammatory responses. The present

XX protein sequence is the human integrin alpha subunit CD11b. This sequence

XX was used in an alignment to identify a novel beta2 integrin alpha

XX subunit: alpha d (AAB07360 and AAB07359). The present sequence has

XX approximately 60% identity to the protein sequence of alpha d. The

XX Alpha d gene and protein may be useful in therapy for diseases linked to

XX sclerosis, asthma, psoriasis, lung inflammation, acute respiratory

XX distress syndrome, rheumatoid arthritis and leukocyte adhesion deficiency

XX (LAD). In addition, anti-alpha d monoclonal antibodies may be used in the

XX inhibition of macrophage infiltration at the site of a central nervous

XX system injury. The monoclonal antibodies can also be used to detect and

XX diagnose Crohn's disease

XX

SQ Sequence 1153 AA;		Query Match 99.13%; Score 5839; DB 3; Length 1153;	
		Best Local Similarity 98.9%; Pred. No. 0;	
		Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;	
QY	1	FNLDTENAMTFQENARGFQSVVQLGSRVVGAPQEI	1080
DB	17	FNLDTENAMTFQENARGFQSVVQLGSRVVGAPQEI	1096
QY	61	RLQVPVEAVNMSLGLSLAATTSPQLLAGCPVHQT	1137
DB	77	RLQVPVEAVNMSLGLSLAATTSPQLLAGCPVHQT	1153
QY	121	FPEALRGCPQSDIAFLIDGSGIIPHDFRRMKELV	1137
DB	137	FPEALRGCPQSDIAFLIDGSGIIPHDFRRMKELV	1153
QY	181	RHFTEKEQNNPNPSLAKPTTOLIGRTHATGLAK	1137
DB	197	RHFTEKEQNNPNPSLAKPTTOLIGRTHATGLAK	1153
QY	241	TGGEKFGDPLGYEDVPIPEADREGVIRYVIGVG	1137
DB	257	TGGEKFGDPLGYEDVPIPEADREGVIRYVIGVG	1153
QY	301	NPEALKTIVQNLREKIPALEGTQTSSTSSFEHEM	1137
DB	317	NPEALKTIVQNLREKIPALEGTQTSSTSSFEHEM	1153
QY	361	GVFLYTSKESKSTFINNTRVDSQNDAYLGYAAI	1137
DB	377	GVFLYTSKESKSTFINNTRVDSQNDAYLGYAAI	1153
QY	421	QNTGMWESNANVKGTOIGAYFGASLCSVDVNSG	1137
DB	437	QNTGMWESNANVKGTOIGAYFGASLCSVDVNSG	1153
QY	481	PRGORARWOCDAVLYGEGQGPWCRFGAALTVDG	1137
DB	497	PRGORARWOCDAVLYGEGQGPWCRFGAALTVDG	1153
QY	541	HGTSGSGISPSHQSRIAGSKLSPRLQYFGQSLG	1137
DB	557	HGTSGSGISPSHQSRIAGSKLSPRLQYFGQSLG	1153
QY	601	PVLVRKAIMFNPVRVARNVFECDQVWKGKEAGE	1137
DB	617	PVLVRKAIMFNPVRVARNVFECDQVWKGKEAGE	1153
QY	661	YDLALDSGRPHSRVAVNETKNSTRROTQVLGLT	1137
DB	677	YDLALDSGRPHSRVAVNETKNSTRROTQVLGLT	1153
QY	721	SLVGTPLSAFGLNLRPVLAEADAQRLFTALPFF	1137
DB	737	SLVGTPLSAFGLNLRPVLAEADAQRLFTALPFF	1153
QY	781	GPREFNVTVVRNDGSDSRTQVTFPPFDLDSYR	1137
DB	797	GPREFNVTVVRNDGSDSRTQVTFPPFDLDSYR	1153
QY	841	SGALKSTCSINHPISFENSEVTFTNITPDVDSK	1137
DB	857	SGALKSTCSINHPISFENSEVTFTNITPDVDSK	1153
QY	901	QLELPVKYAVVWVTSHGVSSTKYLNETASENTR	1137
DB	917	QLELPVKYAVVWVTSHGVSSTKYLNETASENTR	1153
QY	961	RLNQTVINDRPQVTFSENLSSTCHTKERLPSHSD	1137
DB	977	RLNQTVINDRPQVTFSENLSSTCHTKERLPSHSD	1153

QY	1021	FFGIQEEFNATLKGNI	1080
DB	1037	FFGIQEEFNATLKGNI	1096
QY	1081	KVEPEFVNPLPLIVGSSVGGLLALITAAALYKLG	1137
DB	1097	KVEPEFVNPLPLIVGSSVGGLLALITAAALYKLG	1153
RESULT 3			
AAU80252			
ID	AAU80252	standard; protein; 1153 AA.	
XX	AAU80252;		
DT	15-JUL-2002	(first entry)	
XX	Human integrin 1 alpha-M subunit protein.		
DE	Integrin; antiinflammatory; immunosuppression; nephritis; dermatitis;		
XX	inflammatory disease; autoimmune disorder; Crohn's disease;		
KW	human immunodeficiency virus; HIV; myocardial infarction;		
KW	Sjorgen's syndrome; rheumatoid arthritis.		
XX	Homo sapiens.		
OS			
XX	Key	Location/Qualifiers	
PH	Misc-difference 499..500		
FT	/note= "Encoded by GGG CAG AGG"		
XX	WO200218583-A2.		
PN	07-MAR-2002.		
XX	31-AUG-2001; 2001WO-US027227.		
XX	01-SEP-2000; 2000US-0229700P.		
XX	(BLOO-) CENT BLOOD RES INC.		
XX	Springer TA, Shimoaka M, Lu C;		
PI	WPI; 2002-382964/41.		
XX	N-PSDB; ABK50046.		
DR	Modified integrin-I or integrin I-like domain polypeptide useful as an		
XX	immunogen to produce antibodies specific to polypeptide, comprises a		
PT	disulfide bond such that polypeptide is stabilized in a desired		
PT	conformation.		
XX	Disclosure; page 109-112; 112pp; English.		
PS	This invention relates to a modified integrin-I or integrin I-like domain		
XX	polypeptide comprising at least one disulfide bond so that the domain is		
CC	stabilised in a desired conformation. The polypeptide of the invention		
CC	may have antiinflammatory or immunosuppressive activities. The		
CC	polypeptides of the invention have an open conformation and are useful as		
CC	immunogens to produce antibodies that selectively bind to integrin I-		
CC	domain and for identifying a modulator of integrin activity, or of		
CC	interaction of an integrin and a cognate ligand. The polypeptide of the		
CC	invention, or antibodies (preferably anti-LFA-1 antibody) is useful for		
CC	treating or preventing an integrin mediated disorder which is an		
CC	inflammatory or autoimmune disorder in a subject and for inhibiting the		
CC	binding of an integrin to a cognate ligand such as Crohn's disease,		
CC	nephritis; human immunodeficiency virus (HIV), myocardial infarction,		
CC	Sjorgen's syndrome, rheumatoid arthritis, dermatitis. A therapeutic		
CC	composition comprising the peptide of the invention is useful for		
CC	treating an integrin mediated disorder in a subject. The polypeptides		
CC	and/or active or antigenic fragments are useful as reagents for diagnosis		
CC	of integrin-mediated disorders. The present sequence represents the human		
CC	integrin-1 alpha-M protein subunit used to generate the mutant		
CC	polypeptides of the invention		

XX	SQ	Sequence 1153 AA:
		Query Match 99.3%; Score 5839; DB 5; Length 1153;
		Best Local Similarity 98.9%; Pred. No. 0;
		Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;
QY	1	FNLDEANMTFQENARGFGQSVMQLQGSRVVVGAPQEIIVAAANGSLVQCDYSTGSCPEI 60
DB	17	FNLDEVAMTFQENARGFGQSVVLQGSRVVVGAPQELIVAAANGSLVQCDYSTGSCPEI 76
QY	61	RLOVPVEAVNMISLGLSAAATSPQLLACGPTVHTCSENTYVKGLCFPLFSGNLRQQPQK 120
DB	77	RLOVPVEAVNMISLGLSAAATSPQLLACGPTVHTCSENTYVKGLCFPLFSGNLRQQPQK 136
QY	121	PPEARLGCPQEDSDIAFLIDGGSGIIIPHPDRMKELVSTIMEOLKKSKTLPFLMQYSSEEP 180
DB	137	PPEARLGCPQEDSDIAFLIDGGSGIIIPHPDRMKEFVSTWEOQLKKSKTLPFLMQYSSEEP 196
QY	181	RIHPTTFKEQNPNPRSLIKPIITOLLGRTHPTATGLRKVVRELFINTGARKNAFKILPLL 240
DB	197	RIHPTTFKEQNPNPRSLVKEIPTOLLGRTHPTATGIKRVVRELFINTGARKNAFKILVVI 256
QY	241	TGCEKFDPGLGYEVDVIPELDREGVIRYYLFGDAPRSEKSRQEINTVASKPERDHVFQAN 300
DB	257	TGCEKFDPGLGYEVDVIPELDREGVIRYYIVGVGDAPRSEKSRQEINTVASKPERDHVFQVN 316
QY	301	NFEALKTVQNLREKIPFAIBGTQTGSSSFHEHMSQGFSAATNSNGLISTVGSYDWAG 360
DB	317	NFEALKTIQNLREKIPFAIBGTQTGSSSFHEHMSQGFSAATNSNGLISTVGSYDWAG 376
QY	361	GVPLYTSKESKSTFINMTRVDSMDNDAYLVGAALAILNNRVOSVLVGAAPRYQHIGLVAMFR 420
DB	377	GVPLYTSKESKSTFINMTRVDSMDNDAYLVGAALAILNNRVOSVLVGAAPRYQHIGLVAMFR 436
QY	421	QNTGMWESNANVKQTQICAIFFGASLCSDVDVSNGSTDILVIGABHYTEOTRGGOVSVCPL 480
DB	437	QNTGMWESNANVKQTQICAIFFGASLCSDVDVSNGSTDILVIGABHYTEOTRGGOVSVCPL 496
QY	481	PRGORARWCDAVLYGEGOGPWGFGAALTIVLGDVWGDKLTDVAI GAPGEDNRGAVTLF 540
DB	497	PRGORAHWCDAVLYGEGOGPWGFGAALTIVLGDVWGDKLTDVAI GAPGEDNRGAVTLF 556
QY	541	HGTSGSGISPESHRSORIIAGSKLSPLRYFGQSLSGGQDLTMDGLVDLTVGAGHVLLLRSQ 600
DB	557	HGTSGSGISPESHRSORIIAGSKLSPLRYFGQSLSGGQDLTMDGLVDLTVGAGHVLLLRSQ 616
QY	601	PVLRAVKAIMFNPREVARNPFECDNVQVKGKAGEVRVCLHVQKSTDRLRREGQIQSVVT 660
DB	617	PVLRAVKAIMFNPREVARNPFECDNVQVKGKAGEVRVCLHVQKSTDRLRREGQIQSVVT 676
QY	661	YDLALDSGRPHSRVAFNETKNSTRROTQVLGLTOTCETLKLQLPNCIEDPVSPVILRLNF 720
DB	677	YDLALDSGRPHSRVAFNETKNSTRROTQVLGLTOTCETLKLQLPNCIEDPVSPVILRLNF 736
QY	721	SLVGTPLSAFGNLEPVLAEADAQRFLTALPFPEKNCNDNICODDLSITFFSMSLDCLVVG 780
DB	737	SLVGTPLSAFGNLEPVLAEADAQRFLTALPFPEKNCNDNICODDLSITFFSMSLDCLVVG 796
QY	781	GPREFNVTIVRNDDGESYRTQVTFPPPLDISYRKVSTLONQRSQRWRWLACESASTEV 840
DB	797	GPREFNVTIVRNDDGESYRTQVTFPPPLDISYRKVSTLONQRSQRWRWLACESASTEV 856
QY	841	SGALKSTSCSINHIIPFPENSEVTFINIFPDVDSKASLGNKLLKANVTISENNMPRTNKTEF 900
DB	857	SGALKSTSCSINHIIPFPENSEVTFINIFPDVDSKASLGNKLLKANVTISENNMPRTNKTEF 916
QY	901	QLELPVKYAVMVVTSIGVSTKYLIINFITASENTSRWMOHQYQVSNLQORSPIPSLVFLVPPV 960
DB	917	QLELPVKYAVMVVTSIGVSTKYLIINFITASENTSRWMOHQYQVSNLQORSPIPSLVFLVPPV 976
QY	961	RLNQTVIWDREPQVTFSENLSSTCHTKERLPSHSDFLAELRPAPVWNCISIAVCQRIQCIDI 1020

Db	977	RLNQTVWDRPQVTFSENLSTCHTKERLPSHSDFLAELRKA	PVNCISIAVCQRIQCDIP	1036
Qy	1021	FFGIQEFNATLKGNIISFDWYIKTSHNHLIVSTAEILFND	SVFTLLPGQCAFVRSQTET	1080
Db	1037	FFGIQEFNATLKGNIISFDWYIKTSHNHLIVSTAEILFND	SVFTLLPGQCAFVRSQTET	1096
Qy	1081	KVPEFVNPPLPIIVGSSVGLLLALITAALYKLGFFKQYK	DMMSGGPPQGAEPQ	1137
Db	1097	KVPEFVNPPLPIIVGSSVGLLLALITAALYKLGFFKQYK	DMMSGGPPQGAEPQ	1153
RESULT 4				
ABG61469				
ID	ABG61469	standard; protein; 1153 AA.		
AC	AC	ABG61469;		
XX	XX			
DT	27-AUG-2002	(first entry)		
DE		Human Beta2 integrin alphaCD11b subunit.		
XX	XX			
KW	Beta2 integrin; alphaD subunit; CD11c subunit; CD11b subunit; LAD;			
KW	leukocyte adhesion deficiency; inflammatory response; diabetes;			
KW	multiple sclerosis; arthritis; graft atherosclerosis; neuroprotective;			
KW	inflammatory bowel disease; Crohn's disease; ulcerative colitis;			
KW	immune complex alveolitis; leukaemia; ICAM-R; VCAM-R; anti-inflammatory;			
KW	intracellular cell adhesion molecule; vascular cell adhesion molecule;			
KW	locomotor recovery; locomotor damage; locomotor impairment;			
KW	autonomic dysfunction; sensory dysfunction; spinal cord injury.			
OS	Homo sapiens.			
XX	XX			
PN	W02002230380-A2.			
XX	XX			
FD	18-APR-2002.			
XX	XX			
PF	15-OCT-2001; 2001WO-US032059.			
XX	XX			
PR	13-OCT-2000; 2000US-00688307.			
XX	XX			
PA	(ICOS-) ICOS CORP.			
XX	XX			
PI	Gallatin WM, Van Der Vieren M;			
XX	XX			
DR	WPI; 2002-463260/49.			
XX	XX			
PT	Use of an anti-alpha-d monoclonal antibodies for promoting locomotor			
PT	recovery, inhibiting locomotor damage, limiting locomotor impairment, or			
PT	limiting autonomic and sensory dysfunction following spinal cord injury.			
XX	XX			
ES	Example 5; Page 191-194; 270pp; English.			
XX	XX			
CC	The invention relates to promoting locomotor recovery, inhibiting			
CC	locomotor damage, limiting locomotor impairment, or limiting autonomic			
CC	and sensory dysfunction following spinal cord injury by administering an			
CC	anti-alpha d (beta2 integrin alpha2 subunit) monoclonal antibody to a			
CC	spinal cord injury victim. The method also involves the use of a ligand			
CC	selected from ICAM-R or VCAM-1 (intracellular cell adhesion molecule,			
CC	vascular cell adhesion molecule). The method is useful for promoting			
CC	locomotor recovery, inhibiting locomotor damage, limiting locomotor			
CC	impairment, or limiting autonomic and sensory dysfunction following			
CC	spinal cord injury. In particular, the spinal cord injury comprises			
CC	compression of the spinal cord. The antibodies are also useful for			
CC	reducing inflammation at the site of a central nervous system injury. The			
CC	specification also details the identification of Beta2 integrin alphaD			
CC	chinas and proteins, for use in raising the antibodies. Beta2 integrins			
CC	are implicated in diseases such as LAD (leukocyte adhesion deficiency,			
CC	inflammatory response, diabetes, multiple sclerosis, arthritis, graft			
CC	atherosclerosis, inflammatory bowel disease, Crohn's disease, ulcerative			
CC	colitis, immune complex alveolitis and leukaemia. The present sequence is			
CC	a Beta2 integrin alpha subunit sequence included for comparison with the			
CC	Beta2 integrin alphaD protein sequences			
XX	XX			

QY	1021	PFQIQEFNATLKNLSFDWIKYKTSNNHLLIVSTABILFNDVSFTLLPQGGAFVRSQTF	1080
QY	1021	PFQIQEFNATLKNLSFDWIKYKTSNNHLLIVSTABILFNDVSFTLLPQGGAFVRSQTF	1080
Db	1037	PFQIQEFNATLKNLSFDWIKYKTSNNHLLIVSTABILFNDVSFTLLPQGGAFVRSQTF	1096
QY	1081	KVEPPEVPNPLPIVGVSSVGLLILALITAAALYKLGFFFRQYKQKMMSEGGPPCAEPQ	1137
Db	1097	KVEPPEVPNPLPIVGVSSVGLLILALITAAALYKLGFFFRQYKQKMMSEGGPPCAEPQ	1153
RESULT 5			
AAOI14428			
ID	AAOI14428	standard; protein; 1153 AA.	
XX	AAOI14428;		
AC	AAOI14428;		
DT	03-MAY-2002	(first entry)	
XX	Integrin Mac-1 alpha subunit.		
DE	Integrin alpha subunit; variant integrin inserted domain protein;		
KW	Mac-1; integrin alpha subunit; variant integrin inserted domain protein;		
KW	open conformation; integrin related inflammatory disorder;		
KW	integrin related immunological disorder; rheumatoid arthritis; ischaemia;		
KW	refusion; hypovolemic shock; infarction; cerebral shock;		
KW	viral infection; cancer; gene therapy; vaccine;		
KW	bioactive agent screening.		
OS	Unidentified.		
XX	WO200204521-A2.		
PN	17-JAN-2002.		
PD	09-JUL-2001; 2001WO-US021805.		
XX	07-JUL-2000; 2000US-0216600P.		
PR	(CALY) CALIFORNIA INST OF TECHNOLOGY.		
PA	(BLOO-) CENT BLOOD RES.		
PI	Springer T;		
XX	WPI; 2002-148167/19.		
DR	New integrin I domain protein having alteration in at least 2		
XX	noncontiguous regions and exits in an open conformation, useful for		
PT	treating, preventing or suppressing inflammatory or immunological		
PT	disorders.		
XX	Example 1; Fig 1P; 90pp; English.		
PS	The invention comprises structurally biased variant integrin inserted (I)		
XX	domain proteins, wherein the alterations to the protein occur in at least		
CC	two noncontiguous regions. Specifically the variant integrin I domain		
CC	proteins are structurally biased to exist in the open conformation,		
CC	thereby altering the binding ability of the protein. The invention also		
CC	comprises nucleic acids encoding the variant integrin I domain proteins,		
CC	The integrin I domain proteins and nucleic acids are useful for treating,		
CC	preventing or suppressing integrin related inflammatory and immunological		
CC	disorders (e.g. rheumatoid arthritis). The variant integrin I domain		
CC	proteins and nucleic acids can also be used for treating: ischaemia/		
CC	refusion (e.g. hypovolemic shock); infarction; cerebral shock; viral		
CC	infection; and cancer. The variant integrin I domain nucleic acids and		
CC	proteins may be used in gene therapy, as vaccines and to screen for		
CC	bioactive agents. The present amino acid sequence represents the Mac-1		
CC	alpha subunit of integrin		
XX	Sequence 1153 AA;		
SQ			
Query Match	99.3%;	Score 5839;	DB 5; Length 1153;
Best Local Similarity	98.9%;	Pred. No. 0;	
Matches 1124; Conservative	8;	Mismatches	5; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARFGGVSQVVLQGSRRVVGAPQEIIVAAANORGLYQCDYSTGSCBPI 60
DB 17 FNLDTENAMTFQENARFGGVSQVVLQGSRRVVGAPQEIIVAAANORGLYQCDYSTGSCBPI 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHOTCSENYVKGFLGFLGNSLRQOPK 120
DB 77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHOTCSENYVKGFLGFLGNSLRQOPK 136
QY 121 FPALRGCPQEDSDIAPLIDGSGSIIPHPFRMKELVSTIMEOLKSKKTLFSLMQVSEEP 180
DB 137 FPALRGCPQEDSDIAPLIDGSGSIIPHPFRMKELVSTIMEOLKSKKTLFSLMQVSEEP 196
QY 181 RIHFTKEFONNPNRSLKPIITQLLGRTHATGLKVVRELPNINGARKNAFKILFL 240
DB 197 RIHFTKEFONNPNRSLKPIITQLLGRTHATGLKVVRELPNINGARKNAFKILFL 256
QY 241 TDGEKFGDPLGYDVIPELDREGVIRYVLGFGDAPRSEKSRQELNTVASKPRDHVFOAN 300
DB 257 TDGEKFGDPLGYDVIPELDREGVIRYVLGFGDAPRSEKSRQELNTVASKPRDHVFOAN 316
QY 301 NFPAKATVQNLREKIFALEGTCTGSSSPHEHMSQEGSAATTSNGPLLLSTVGSVDWAG 360
DB 317 NFPAKATVQNLREKIFALEGTCTGSSSPHEHMSQEGSAATTSNGPLLLSTVGSVDWAG 376
QY 361 GVFLYTSKEKSTFINNTRVDSNDNDAYLGAAAIILNRNRVQSLVLCAPRYQHIGLVAMPR 420
DB 377 GVFLYTSKEKSTFINNTRVDSNDNDAYLGAAAIILNRNRVQSLVLCAPRYQHIGLVAMPR 436
QY 421 QNTGMESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYVTEGQSVCP 480
DB 437 QNTGMESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYVTEGQSVCP 496
QY 481 PRQORARWQCDVLYGEOGQPMGRFGAALTVLGDVANGDKLTVAIAGPGEEDNRGAVYLF 540
DB 497 PRQORARWQCDVLYGEOGQPMGRFGAALTVLGDVANGDKLTVAIAGPGEEDNRGAVYLF 556
QY 541 HGTSGSGISPSHQRAGSKSLPRLQYFGQSLSGGQDLTWGGLVDLTVGAQGHVLLRSQ 600
DB 557 HGTSGSGISPSHQRAGSKSLPRLQYFGQSLSGGQDLTWGGLVDLTVGAQGHVLLRSQ 616
QY 601 PVLKVKALMEFNRVAVNVEFNCNDQVKGKEAGEVRVCLHVQKSTRDLRREGQIQSVVT 660
DB 617 PVLKVKALMEFNRVAVNVEFNCNDQVKGKEAGEVRVCLHVQKSTRDLRREGQIQSVVT 676
QY 661 YDLALDSGRPHSRVAVNETKSTRQTVLGLTQCTETLKLQLPNCIEDPVPSPVILRNF 720
DB 677 YDLALDSGRPHSRVAVNETKSTRQTVLGLTQCTETLKLQLPNCIEDPVPSPVILRNF 736
QY 721 SLVGTPLSAFQNLRPVLAEDAQLFTALPPEKNCNDNLCQDDLSITFSFMSLDCLVG 780
DB 737 SLVGTPLSAFQNLRPVLAEDAQLFTALPPEKNCNDNLCQDDLSITFSFMSLDCLVG 796
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSQRSLRACASSTEV 840
DB 797 GPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSQRSLRACASSTEV 856
QY 841 SGALKSTCSINHPITPENSEVFNITFDVDSKASLGNKLLKANVTSENNWERTKTEP 900
DB 857 SGALKSTCSINHPITPENSEVFNITFDVDSKASLGNKLLKANVTSENNWERTKTEP 916
QY 901 QLELPVKYAVTMVTSHGVSSTKYNLFTASENTSRVMOHQYQVSNLQORSLPISLVLPV 960
DB 917 QLELPVKYAVTMVTSHGVSSTKYNLFTASENTSRVMOHQYQVSNLQORSLPISLVLPV 976
QY 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELKAPVNCSTIACVORICDIP 1020
DB 977 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELKAPVNCSTIACVORICDIP 1036
QY 1021 FFGQIESFNATLKNLSFDMYKTSNHLIIVSTAELFNDVSFTLLPGQGFVRSQTET 1080
DB 1037 FFGQIESFNATLKNLSFDMYKTSNHLIIVSTAELFNDVSFTLLPGQGFVRSQTET 1096
QY 1081 KVPEFVENPPLIVGSSVGGLLLLALITAAALYKLGFFKQYKDMMSGGPPGABEQ 1137

DB 1097 KVPEFVENPPLIVGSSVGGLLLLALITAAALYKLGFFKQYKDMMSGGPPGABEQ 1153
RESULT 6
ADD25615
ID ADD25615 standard; protein; 1153 AA.
XX
AC ADD25615;
DT 15-JAN-2004 (first entry)
XX
DE Binding domain-immunoglobulin fusion protein-associated protein #85.
DB
DE Binding domain; immunoglobulin; fusion protein; cytostatic;
antiarrhythmic; immunosuppressive; antidiabetic; antichryoid;
neuroprotective; hinge region; immunoglobulin heavy chain;
KW CH2 constant region; CH3 constant region; IgG1;
KW antibody dependent cell-mediated cytotoxicity; ADCC; complement fixation;
KW rheumatoid arthritis; B-cell disorder; melanoma; carcinoma;
KW rheumatoid arthritis; myasthenia gravis; Grave's disease;
KW type I diabetes mellitus; multiple sclerosis; autoimmune disease.
XX
OS Unidentified.

US2003118592-A1.

26-JUN-2003.

25-JUL-2002; 2002US-00207655.

17-JAN-2001; 2001US-0367358P.

17-JAN-2002; 2002US-00053530.

03-JUN-2002; 2002US-0385691P.

(GENE-) GENE-CRAFT INC.

Ledbetter JA, Hayden-Ledbetter MS, Thompson PA;

WPI; 2003-801317/75.

New binding domain-immunoglobulin fusion protein, useful for treating a subject having or suspected of having a malignant condition or a B-cell disorder, e.g. melanoma, Grave's disease or autoimmune disease.

Disclosure; SEQ ID NO 176; 157pp; English.

The invention relates to a binding domain-immunoglobulin fusion protein comprising a binding domain polypeptide that is fused to an immunoglobulin hinge region polypeptide, an immunoglobulin heavy chain CH2 constant region polypeptide that is fused to the hinge region polypeptide, and an immunoglobulin heavy chain CH3 constant region polypeptide that is fused to the CH2 constant region polypeptide. The hinge region polypeptide comprises: a wild-type human IgG1 immunoglobulin hinge region polypeptide; a mutated human IgG1 immunoglobulin hinge region polypeptide, derived from (a) having 3 or more cysteine residues; where the mutated human IgG1 immunoglobulin hinge region polypeptide contains 2 cysteine residues, where the first cysteine is not mutated; a mutated human IgG1 immunoglobulin hinge region polypeptide, derived from (a) having 3 or more cysteine residues, where the mutated human IgG1 immunoglobulin hinge region polypeptide contains no cysteine residues. The binding domain-immunoglobulin fusion protein is capable of at least one immunological activity comprising antibody dependent cell-mediated cytotoxicity (ADCC) and complement fixation. The binding domain polypeptide is capable of specifically binding to an antigen. Also included are an isolated polynucleotide encoding the binding domain-immunoglobulin fusion protein, a recombinant expression construct comprising the polynucleotide operably linked to a promoter, a host cell transformed or transfected with a recombinant expression construct, producing the binding domain-immunoglobulin fusion protein, a

pharmaceutical composition comprising the binding domain-immunoglobulin fusion protein or polynucleotide and a carrier, and treating a subject having or suspected of having a malignant condition or a B-cell disorder. The binding domain-immunoglobulin fusion protein is useful for treating a subject having or suspected of having a malignant condition or a B-cell disorder, e.g. melanoma, carcinoma or sarcoma, rheumatoid arthritis, myasthenia gravis, Grave's disease, type I diabetes mellitus, multiple sclerosis or autoimmune disease. The present sequence is a binding domain -immunoglobulin fusion protein-associated protein sequence. Note: The sequence data for this patent formed part of the printed specification and is also available in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html?docID=20030118592. The authors have not identified the sequences in the printed specification by their SEQ ID number therefore none of the sequences can be explicitly identified.

Query Match	99.3%	Score 5839;	DB 7;	Length 1153;
Best Local Similarity	98.9%	Pred. NO. 0;		
Matches 1124; Conservative		8; Mismatches	5; Indels	0; Gaps

Qy	1	FNLDTENAMTFQENARGFGQSVQVLQSGRVVVGAPQEIIVAAORGSLLVQCDYSTGSCBPI	60
Db	17	FNLDTENAMTFQENARGFGQSVQVLQSGRVVVGAPQEIIVAAORGSLLVQCDYSTGSCBPI	76
Qy	61	RLOVPVEAVNMSLGLSLAATTSFPOLLACGPTVHOTCSENTYVXGLCLFSGNLNRQOQXK	120
Db	77	RLOVPVEAVNMSLGLSLAATTSFPOLLACGPTVHOTCSENTYVXGLCLFSGNLNRQOQXK	136
Qy	121	FPEALRGCPQSDSDIAFLIDSGSGIIIPHDFFRMKELVSTIMEOLUKSKXTLLFSLMOYSBEF	180
Db	137	FPEALRGCPQSDSDIAFLIDSGSGIIIPHDFFRMKEFVSTVMEOLUKSKXTLLFSLMOYSBEF	196
Qy	181	RIHETFEKBFQNNPFRSLIKPITOLLGRTHATGLRKVVURELENTITNGARKNAFKILL	240
Db	197	RIHETFEKBFQNNPFRSLIKPITOLLGRTHATGLRKVVURELENTITNGARKNAFKILLVVI	256
Qy	241	TDGEKFGDPLGYEYDVIPELBDREGVIRYVVGFDAPRSEKSKOEINTVASKPPRDHVFOAN	300
Db	257	TDGEKFGDPLGYEYDVIPEADREGVIRYVVGFDAPRSEKSKOEINTIASKPPRDHVFOVN	316
Qy	301	NFEALKTVQNLREKIFALEICTQTGSSSSFEHMSQEGFSAAITSNGLPSTVSGSYDNAG	360
Db	317	NFEALKTIQNLREKIFALEICTQTGSSSSFEHMSQEGFSAAITSNGLPSTVSGSYDNAG	376
Qy	361	GVFLYTSKEKSTFTINMTRVDSMDNDAYLGYAAAIIILNRVQSLVLGAPRYOHIGLVAMFR	420
Db	377	GVFLYTSKEKSTFTINMTRVDSMDNDAYLGYAAAIIILNRVQSLVLGAPRYOHIGLVAMFR	436
Qy	421	QNTGWSNANVKGTQIGAYFGASILCSVDVDSNGSTDVLVIGAPHYBQTEGGQSVVCPL	480
Db	437	QNTGWSNANVKGTQIGAYFGASILCSVDVDSNGSTDVLVIGAPHYBQTEGGQSVVCPL	496
Qy	481	PRGORARWQCDAVLYGBQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLP	540
Db	497	PRGORARWQCDAVLYGBQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLP	556
Qy	541	HGTSGSGISPSHSORIASGKSLSPRIQYFGQSLSGSQDLTMDGLVDLTVGAQGHVLLRSQ	600
Db	557	HGTSGSGISPSHSORIASGKSLSPRIQYFGQSLSGSQDLTMDGLVDLTVGAQGHVLLRSQ	616
Qy	601	PVLRVKAIMFNPREVARNVFECDQVQVKGKEAGEVRVCLHVQKSTRDLRLEGQIQSVVT	660
Db	617	PVLRVKAIMFNPREVARNVFECDQVQVKGKEAGEVRVCLHVQKSTRDLRLEGQIQSVVT	676
Qy	661	YDLALDSGRPHSRVAFNETKSTRTOVLGLTQCEFLKQLPNCIEDPVSPIVLRNLF	720
Db	677	YDLALDSGRPHSRVAFNETKSTRTOVLGLTQCEFLKQLPNCIEDPVSPIVLRNLF	736
Qy	721	SLVGTPLSAFNLRPVLAEDAQRLFTALFPPEKNCNDNI CODDLSITFFSMSLDCLVVG	780
Db	737	SLVGTPLSAFNLRPVLAEDAQRLFTALFPPEKNCNDNI CODDLSITFFSMSLDCLVVG	796

Qy	781	GPREFNVTVTRNDGSDSYRTQTVPFPFPLDLSTRKVTSTTLQNQSORSQRWLACESASSTEY	840
Dd	797	GPREFNVTVTRNDGSDSYRTQTVPFPFPLDLSTRKVTSTTLQNQSORSQRWLACESASSTEY	856
Qy	841	SGALKSTCSINHIPIIPENSEVFNITTFDVSXASLGNKLLLKANTVTSENMMPTNKTEP	900
Dd	857	SGALKSTCSINHIPIIPENSEVFNITTFDVSXASLGNKLLLKANTVTSENMMPTNKTEP	916
Qy	901	QLLEPVKYAYVMVTVTSHGVSTKYLNFTAGENTSRRVMQHGYQVNGLGORSILPISLVFLVPV	960
Dd	917	QLLEPVKYAYVMVTVTSHGVSTKYLNFTAGENTSRRVMQHGYQVNGLGORSILPISLVFLVPV	976
Qy	961	RLNQTVTWDRPOVTFSENTSSTCHTKERLPESHDSDELAEELRKAIVMNCIAVCORIOCDIP	1020
Dd	977	RLNQTVTWDRPOVTFSENTSSTCHTKERLPESHDSDELAEELRKAIVMNCIAVCORIOCDIP	1036
Qy	1021	FFGIQSEFNATLKGNIISFDWIYIKTSHNNLLIVTAIELFNDSYFTLLPGQGAFVRSGTOT	1080
Dd	1037	FFGIQSEFNATLKGNIISFDWIYIKTSHNNLLIVTAIELFNDSYFTLLPGQGAFVRSGTOT	1096
Qy	1081	KVPEPFENVNPLPLIVSSVCGILLIALLITAAALYKLGFKKROYKDMMSGPPGAEPQ	1137
Dd	1097	KVPEPFENVNPLPLIVSSVCGILLIALLITAAALYKLGFKKROYKDMMSGPPGAEPQ	1153
 RESULT 7			
AAR04136			
ID	AAR04136	standard; protein; 1153 AA.	
XX			
AC	AAR04136;		
XX			
DT	25-MAR-2003	(revised)	
DT	07-SEP-1990	(first entry)	
XX			
DE	Alpha subunit of Mac-1 leukocyte adhesion receptor.		
XX			
KW	Mac-1 alpha subunit; Mac-1 alpha/beta heterodimer;		
XW	non-specific defence system; integrin gene superfamily.		
XX			
OS	Synthetic.		
XX			
FH	Key	Location/Qualifiers	
FT	Region	1..116	
FT	Modified-site	/label= signal peptide	
FT	Modified-site	/label= putative N-glycosylation site	
FT	Modified-site	240..242	
FT	Modified-site	/label= putative N-glycosylation site	
FT	Modified-site	391..393	
FT	Modified-site	/label= putative N-glycosylation site	
FT	Modified-site	469..471	
FT	Modified-site	/label= putative N-glycosylation site	
FT	Modified-site	693..695	
FT	Modified-site	/label= putative N-glycosylation site	
FT	Modified-site	697..699	
FT	Modified-site	/label= putative N-glycosylation site	
FT	Modified-site	735..737	
FT	Modified-site	/label= putative N-glycosylation site	
FT	Modified-site	802..804	
FT	Modified-site	/label= putative N-glycosylation site	
FT	Modified-site	881..883	
FT	Modified-site	/label= putative N-glycosylation site	
FT	Modified-site	901..903	
FT	Modified-site	/label= putative N-glycosylation site	
FT	Modified-site	912..914	
FT	Modified-site	/label= putative N-glycosylation site	
FT	Modified-site	941..943	
FT	Modified-site	/label= putative N-glycosylation site	
FT	Modified-site	947..949	
FT	Modified-site	/label= putative N-glycosylation site	
FT	Modified-site	979..981	
FT	Modified-site	/label= putative N-glycosylation site	
FT	Modified-site	994..996	

FT	Modified-site	/label= putative N-glycosylation site	
FT	Modified-site	1022..1024	
FT	Modified-site	/label= putative N-glycosylation site	
FT	Modified-site	1045..1047	
FT	Modified-site	/label= putative N-glycosylation site	
FT	Modified-site	1051..1053	
FT	Modified-site	/label= putative N-glycosylation site	
FT	Modified-site	1076..1078	
FT	Modified-site	/label= putative N-glycosylation site	
FT	Modified-site	1106..1134	
FT	Region	/label= putative transmembrane region	
XX		EP364690-A.	
XX		25-APR-1990.	
XX		17-AUG-1989;	89EP-00115159.
XX		23-AUG-1988;	89US-00235353.
XX		09-MAR-1989;	89US-00321239.
XX		(DAND) DANA FARBER CANCER INST INC.	
XX		Springer TA, Corbi A;	
XX		WPI; 1990-125938/17.	
XX		N-ESDB; AAQ04043.	
XX		New pure Mac-1 alpha sub-unit and corresp. DNA - useful for treating	
XX		inflammation and viral infections, and in diagnosis.	
XX		Disclosure; Page ?; -pp; English.	
XX		Mac-1 alpha subunit is involved in the response to inflammation, i.e.	
XX		recognition of and migration to sites of inflammation. It also attaches	
XX		to cellular substrates as part of this function making it useful in	
XX		visualising endothelial tissue. Mac-1 is a member of the Integrin Gene	
XX		superfamily. (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25	
XX		-MAR-2003 to correct PA field.)	
XX		Sequence 1153 AA:	
XX		Query Match	99.1%; Score 5829; DB 2; Length 1153;
XX		Best Local Similarity	98.8%; Pred. No. 0;
XX		Matches 1123; Conservative	8; Mismatches 5; Indels 0; Gaps 0;
QY		1 FNLDTENAMTFOENARGGQSVVQLQGSRRVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI	60
DB		17 FNLDTENAMTFOENARGGQSVVQLQGSRRVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI	76
QY		61 RLQVPVEAVNLSGLSLAATTSPOLLACGPTVHQTCSNTYVVGKGLCFPGSNLRQPOK	120
DB		77 RLQVPVEAVNLSGLSLAATTSPOLLACGPTVHQTCSNTYVVGKGLCFPGSNLRQPOK	136
QY		121 PPEALRGCPQSDSDIAFLIDGSGTIIPHDPRMKELVSTIMEQLKSKTLPFLMQYSEEF	180
DB		137 PPEALRGCPQSDSDIAFLIDGSGTIIPHDPRMKELVSTIMEQLKSKTLPFLMQYSEEF	196
QY		181 RIHPTFKFQNNPRLSKITLQGLTHTATGLKVVRELVNITGARKNAKILFL	240
DB		197 RIHPTFKFQNNPRLSKITLQGLTHTATGLKVVRELVNITGARKNAKILVVI	256
QY		241 TDGEKFGDPLGYEDVPEADREGVIRYVVGDAFRSEKSRQELNTVASKPRPDHVFQAN	300
DB		257 TDGEKFGDPLGYEDVPEADREGVIRYVVGDAFRSEKSRQELNTVASKPRPDHVFQV	316
QY		301 NFEALKTQNLREKIPALEGTQTCSSSPHEMSQEGFSAITNSGELLSTGSDYDAG	360
DB		317 NFEALKTQNLREKIPALEGTQTCSSSPHEMSQEGFSAITNSGELLSTGSDYDAG	376
QY		361 GVFLYTSKEKSTFFINWTVDSMDNDAYLGAAAIILNRVQSLVGLGAPRYOHIGLVAMFR	420
DB		377 GVFLYTSKEKSTFFINWTVDSMDNDAYLGAAAIILNRVQSLVGLGAPRYOHIGLVAMFR	436

Key	Location/Qualifiers
Region	1..19
Region	/label= signal peptide
Region	20..44
Modified-site	/label= N-terminus
Modified-site	61..63
Modified-site	/label= glycosylation site

p150.95 leucocyte adhesion receptor alpha-subunit; hairy cell leukaemia; rhinovirus.
Synthetic.
p150.95 alpha subunit encoded by clone lambdaX47.

RESULT 8
AAR07120
ID AAR07120 standard; protein; 1163 AA.
AC AAR07120;
XX
DT 25-MAR-2003 (revised)
DT 05-FEB-1991 (first entry)
XX
DB p150.95
XX
KW p150.95
XX
OS Synthetic.

1081 KVEPFVVPPLLIIVGSSVGGILLALITAAALYKLGPFKXOYKDMSEGGPPCAEPQ 1137
1097 KVEPFVVPPLLIIVGSSVGGILLALITAAALYKLGPFKXOYKDMSEGGPPCAEPQ 1153

1021 PFGIQEFPNATLKGNSLFDWYIKTSHNLLIVSTABILFNDVSFTLLPGQAFVRSQTET 1080
1037 PFGIQEFPNATLKGNSLFDWYIKTSHNLLIVSTABILFNDVSFTLLPGQAFVRSQTET 1096

961 RLNOTVINDRPQVTFSENLSSTCHTKERLPSHSDFLAEARKAPVNCSTIACVQRIQCDIP 1020
977 RLNOTVINDRPQVTFSENLSSTCHTKERLPSHSDFLAEARKAPVNCSTIACVQRIQCDIP 1036

901 QLELPVKYAVYMWVTSYKYNFTASENTSRVMQHOYQVSNLQORSLPISLVLPV 960
917 QLELPVKYAVYMWVTSYKYNFTASENTSRVMQHOYQVSNLQORSLPISLVLPV 976

841 SGALKSTCSINHPIFPENSEVTFNITFDVDSKASLGKLLKANKVTSNNMPRTNKTEF 900
857 SGALKSTCSINHPIFPENSEVTFNITFDVDSKASLGKLLKANKVTSNNMPRTNKTEF 916

781 GPREFNVTVNRDGEDSYRTQVTFPPDLDSYRKVSTLONQORSQSWLACESASSTEV 840
797 GPREFNVTVNRDGEDSYRTQVTFPPDLDSYRKVSTLONQORSQSWLACESASSTEV 856

721 SLVGTPLSAFAGNLRPVLAEDAQLFTALFPFEKNCNDNICODDLSITTFMSLDCIIVG 780
737 SLVGTPLSAFAGNLRPVLAEDAQLFTALFPFEKNCNDNICODDLSITTFMSLDCIIVG 796

661 YDLALDSGRPHSRAPVNETKNSRTQTQVGLTQTCETLKLQLPNCIEPVPVILRLNF 720
677 YDLALDSGRPHSRAPVNETKNSRTQTQVGLTQTCETLKLQLPNCIEPVPVILRLNF 736

601 PVLRVKAIIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLRLBQSIQSVVT 660
617 PVLRVKAIIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLRLBQSIQSVVT 676

541 HGTSGSGISPSHSQRIAGSKLSPRLQYFQGSISGGODLTMDGLVLTVCAGQGHVLLRSQ 600
557 HGTSGSGISPSHSQRIAGSKLSPRLQYFQGSISGGODLTMDGLVLTVCAGQGHVLLRSQ 616

481 PRGORARWOCDAVLYGEGQCPWGRFGAALTVDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
497 PRGORARWOCDAVLYGEGQCPWGRFGAALTVDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

437 QNTGMWESNANVKGTQIGAYFGASICSVDVDSNGSTDVLVIGAPHYETOTRGQSVVCP 480
437 QNTGMWESNANVKGTQIGAYFGASICSVDVDSNGSTDVLVIGAPHYETOTRGQSVVCP 496

421 QNTGMWESNANVKGTQIGAYFGASICSVDVDSNGSTDVLVIGAPHYETOTRGQSVVCP 480
437 QNTGMWESNANVKGTQIGAYFGASICSVDVDSNGSTDVLVIGAPHYETOTRGQSVVCP 496

KW Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;
 KW type-I diabetes; atherosclerosis; multiple sclerosis; asthma;
 KW lung inflammation; acute respiratory distress syndrome; CD11c subunit;
 KW rheumatoid arthritis.

XX Homo sapiens.

XX US5728533-A.

XX 17-MAR-1998.

XX 07-JUN-1995; 95US-00485618.

XX 23-DEC-1993; 93US-00173497.

XX 05-AUG-1994; 94US-00286889.

XX 21-DEC-1994; 94US-00362652.

XX (ICOS-) ICOS CORP.

XX Van Der Vieren M, Gallatin WM;

XX W21; 1998-206565/18.

XX Screening assay for modulators of integrin binding - using immobilised or
 labelled alpha-d polypeptide, useful for, e.g. treating type-I diabetes.

XX Example 5; Fig 1A-D; 106pp; English.

XX This sequence represents a human beta-integrin CD11c subunit which is
 used to describe a method for identifying compounds that modulate the
 interaction of the beta-integrin alpha-d subunit with a binding partner
 of alpha-d which involves contacting an alpha-d polypeptide with an alpha
 -d binding partner, one of which is immobilised and the other of which is
 labelled, in the presence of a test compound, and determining if the
 compound affects binding between the alpha-d polypeptide and alpha-d
 binding partner, where the alpha-d polypeptide is alpha-d or its fragment
 comprising the cytoplasmic, transmembrane or extracellular domain of
 alpha-d. Compounds that modulate alpha-d binding could be used to treat
 diseases such as type-I diabetes, atherosclerosis, multiple sclerosis,
 asthma, psoriasis, lung inflammation, acute respiratory distress syndrome
 and rheumatoid arthritis

XX Sequence 1163 AA;

Query Match 58.2%; Score 3423; DB 2; Length 1163;

Best Local Similarity 60.2%; Pred. No. 2.7e-274;

Matches 680; Conservative 149; Mismatches 294; Indels 6; Gaps 4;

1 FNLDTNMTAFQNRAGPGQSVVQLQGSVVVVGAPQEIIVAAVNRGSLVQCDYSTGSCPEI 60

20 FNLDTNMTAFQNRAGPGQSVVQLQGSVVVVGAPQEIIVAAVNRGSLVQCDYSTGSCPEI 79

61 RLQVPEAVNMSLGLSIATTPSPQLLACGPTVHQCSTENTYVKGCLFQSNLRQOPQX 120

80 GLQVPEAVNMSLGLSIATTPSPQLLACGPTVHQCSTENTYVKGCLFQSNLRQOPQX 137

121 FPLALGCCPQEDDIAFLDGGSIIPHPFMKSELVSTIMEQLKSKTFLSLMQYSEEF 180

138 LPVSRQCEQEDIVFLDGGSIIPHPFMKSELVSTIMEQLKSKTFLSLMQYSEEF 197

181 RHFTTKFQNNPNSLTKPTTOLLGRTHATGRKURVRELNITGACRQAKLIFLL 240

198 QHTFTFEEFRNTSNPLUSLASVHQLQGFYTTATQNVVHRLPHASYGARDAKILIVI 257

241 TDGEKFGDPLGVEDVPELDRBGVIRYVLGFGDAFSEKSRQELMTVASKEPRDHFQAN 300

258 TDGKKEGSDLYKDVIPMDAAGIIRYVGLAFONRNSNKLNDIAKSPQSHIFKVE 317

301 NPEALXTQVQNRKELPALEGTTQTSQSSSEHEMSQEGFSAITSNGLLSITGVSQWAG 360

318 DFDALNDIQNLKELPALEGTTQTSQSSSEHEMSQEGFSAITSNGLLSITGVSQWAG 377

361 GVFLYTSKEKSTFINNTRYDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYOHIGLVAMFR 420

378 GAFLYPPNMSPTFTNMQENVDMDSDYLGSTELALWKGVSQSLVGLGAPRYOHIGKAVIFI 437
 421 QVTKGWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQFGQGVSCPL 480
 438 QVSRQWRMKAQVIGTQIGSYFGASLCSVDVDSNGSTDLVLIGAPHYYEQFGQGVSCPL 497
 481 PRGQARHQCDAVLYGEGQFAGALTVLGDVAGDKLTVLVAIGAPGEDHREGAVYLP 540
 498 PRGMR-RWCCDAVLYGEGQFAGALTVLGDVAGDKLTVLVAIGAPGEDHREGAVYLP 556
 541 HCTSGSISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVGAGQHVLLRSQ 600
 557 HGVLPSPISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVGAGQHVLLRSQ 616
 601 PVLRYKATMBENPREVAENVECNQVVKGEAGEVCLHVOKSTFDRLEQIOISVVT 660
 617 PVLWVGVSQNFIPAEIPSAFECRQVVSQTLVQSNICLIYIDRKNLLGSDLSVVT 676
 661 YDLALDSCRPHSRVAFNETQNSTROTQVIGLGTOTCETLKLQLPNCHIEDVPSIVLRNF 720
 677 LDALAPGRLSPRAIFQETKRSLSRVRLGLKAHCENFNLLPSCVEDSVIPIRLNF 736
 721 SLVGTPLSAGNLRPVLAEDAQRFTALPFPEKNCNDNICODDLSITPFSMLDCLVVG 780
 737 TLVGPPLAFNLRPMLAALAQRFTASLPFEKNCADHICODNLSISFSPGLKSLVG 796
 781 GPRENVTVTVNRDGEDSYRTQVTFPFDLSYKRVSTLQNRQSRQSRWRLACESASTEV 840
 797 SNLELNAMVWVNDGEDSYRTQVTFPFDLSYKRVSTLQNRQSRQSRWRLACESASTEV 854
 841 SCALKSTSCSINHPIPPENSEVITNITFDVDSKASLGNKLLKXANTVSENMPRTWKTEF 900
 855 SQGTWSTSCRINELIFRCGAQITLAFDVSRAVGLDRLLLIANVSSENIPRTSKTIF 914
 901 QLELPYKAVYVTVSHCVSKYLNFTAS-ENTSRVMOHOYVSNLQSRSLVPLVP 959
 915 QLELPYKAVYVTVSHCVSKYLNFTAS-ENTSRVMOHOYVSNLQSRSLVPLVP 974
 960 VELNQTIVDWPQVTFSENLSSTCHYKRLPSSHSDFLAELKAPVNVNCISAVCQIQCDI 1019
 975 VELNQTIVDWPQVTFSENLSSTCHYKRLPSSHSDFLAELKAPVNVNCISAVCQIQCDI 1034
 1020 PFGIOEEFNATLKNLSFDWYIKTSHNHLIYVSTAEILFNDVSFTLLPCGGAFVRSQTE 1079
 1035 PSFVSQESLDTLKNLSFGVWRQILQKKVSVSVASIIIFDTSVYSOLPQGEAFMRAQTI 1094
 1080 TKVEPFYVNPPLIVGSSVGLLLALITAAALYKLGFFKQYKQWMSSE 1128
 1095 TVLEKYKVHNPILVIGSSIGGLLLALITAAALYKLGFFKQYKQWMSSE 1143

RESULT 10
 AAB07361
 ID AAB07361 standard; protein; 1163 AA.
 XX
 AC AAB07361;
 DT 17-JAN-2001 (first entry)
 XX
 DE Human CD11c protein sequence.
 XX
 KW Human; macrophage infiltration inhibition; alpha_d integrin;
 KW leukocyte integrin; Leu-CAM; leukocyte integrin; immune response;
 KW inflammation; leukocyte adhesion deficiency; LAD; Type I diabetes;
 KW atherosclerosis; multiple sclerosis; asthma; psoriasis; beta2 integrin;
 KW lung inflammation; acute respiratory distress syndrome; Crohn's disease;
 KW rheumatoid arthritis; central nervous system injury; CD11c.
 XX
 OS Homo sapiens.
 XX
 XX WO200029446-A1.
 XX

FT	Region	/note= "extracellular domain"	
FT	150..352		
FT	/note= "this region is homologous to the insert common to Chla,b,c and may be a site for interaction with ICAM family proteins"		
FT	465..474		
FT	Binding-site	/note= "putative cation binding site"	
FT	518..527		
FT	Binding-site	/note= "putative cation binding site"	
FT	592..600		
FT	Binding-site	/note= "putative cation binding site"	
FT	1109..1128		
FT	Region	/note= "transmembrane region"	
FT	1129..1161		
FT	Domain	/note= "cytoplasmic domain"	
XX	WO9517412-A1.		
XX	29-JUN-1995.		
XX	21-DEC-1994;	94WO-US014832.	
XX	23-DEC-1993;	93US-00173497.	
XX	05-AUG-1994;	94US-00286889.	
XX	(ICOS-) ICOS CORP.		
XX	Gallatin WM, Van Der Vieren M;		
XX	WPI; 1995-240603/31.		
XX	N-PSDB; AAQ91712.		
XX	Alpha sub-unit polypeptide of human beta 2 integrin - used to identify potential antiinflammatory agents, for the treatment of graft arteriosclerosis, inflammatory bowel disease, asthma, etc.		
XX	Claim 7; Page 82-87; 172pp; English.		
XX	A probe based on a partial cDNA clone (given in AAQ91727) of canine alpha -TM1 was used to screen a human spleen cDNA library to identify clone 19A2 encoding the beta-1 integrin alpha-subunit, alpha-d. The cDNA was manipulated to allow expression of recombinant alpha-d subunit in COS and CHO cells		
XX	Sequence 1161 AA;		
XX	Query Match	57.6%; Score 3388; DB 2; Length 1161;	
XX	Best Local Similarity	58.9%; Pred. No. 2.2e-271;	
XX	Matches	665; Conservative 168; Mismatches 288; Indels 8; Gaps 6;	
Qy	1	FNLDTENAMFQENARGFGQSVVQLOGSRVVGAPQEIIVAAQSGSLYQCDYSTGSCPEI 60	
Db	17	FNLDEEFTTFQEDAGFGQSVVQFGSRVLVVGAPLEVAANOTGRLYDCAATGMCQPI 76	
Qy	61	RLQVPVEAVNMSGLSLAATSPQLACGPTVHQTCSENTYVVKGLCFLFGSNLRQOPQK 120	
Db	77	PLHIRPEAVNMSGLTTLAATNGSRLACGPTLHRVCGENSYSKSGCLLGSRW-BIIQT 135	
Qy	121	PPALRCQEDSDIARLIDGSGIIIPDRFRRKELVSTIMEQLKSKTLPISLMQYSEEF 180	
Db	136	VPDAPTECPHQEMDIVFLIDGSGIDQNDENQMKFVQAVMGQFEGDTLPAFMQYSNLL 195	
Qy	181	RIHFTFKFQNNPRLIKPITOLLGRTHATGLRVRVRELFTNITNGARKNAFKILELL 240	
Db	196	KIHFTTFQFTSPSQSLVDPVQLKGLTTFATGILTVTVQLFHHKNGARKSAKKILIVI 255	
Qy	241	TGDKFGDPIGYENVIPELDREGVIRVLVFGDAPFSEKQELNIVASPPDDHVPQAN 300	
Db	256	TGQKYDPLFYSVDPQAEAGIIRYAIGVGHAFQGPQTARQELNITISSAPPQDHVKVD 315	
Qy	301	NFEALKTVQNLREKI PAIEGTGTGSSSSEHEMSQEGPSAAITSNGLLSTVGSVDWAG 360	
Db	316	NFAALGSIQLOKEIVAVGTSQRASSSFQHEMSQEGFTALTMDGLFGAVGSFWSG 375	
Qy	361	GVFLYTSKEKSTPIFNMTRVDSMDNDAYLGIAAAIILRNVRQSLVLCAPRYQHILVAMPR 420	
Db	376	GAPLPNNKSPPIFNMNSQENVRDSDYLGISTELAKWGVQNLVLCAPRYQHTKAVIFT 435	
Qy	421	QNTGWESNANVKGTQIGAYFGASLCSDVDNSGSDTDLVLICAPHYEYQTRGQSVCPFL 480	
Db	436	QVSRQWRKKAETVGTQIGSYFGASLCSDVDSDGSDTDLILICAPHYEYQTRGQSVCPFL 495	
Qy	481	PRGQARWOCDAVLYGBOGQPWGRFGAALTVDVNGDKLTDVAIGAPGEENRGAVYLP 540	
Db	496	PRGQARWOCDAVLYGBOGQPWGRFGAALTVDVNGDKLTDVAIGAPGEENRGAVYLP 555	
Qy	541	HGTSGGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVLTGAOQHLLLRQ 600	
Db	556	HGASESGISPSHSQRIASSQLSPRLQYFGQSLSGGQDLTMDGLVLTGAOQHLLLRQ 615	
Qy	601	PVLRYKAIMFNPVRVARNVFCNDQVVKGEAGVRYVCLHVQKSTRRLRREGQTSVVT 660	
Db	616	PVLRYKAIMFNPVRVARNVFCNDQVVKGEAGVRYVCLHVQKSTRRLRREGQTSVVT 673	
Qy	661	YDLALDSGRPHSRVFNENSTRTQTOVLGTQTCETLKLQLPNCIEDPVPVILRNPF 720	
Db	674	YDLALDSGRPHSRVFNENSTRTQTOVLGTQTCETLKLQLPNCIEDPVPVILRNPF 733	
Qy	721	SLVGTPLSAGNLRPVLAEQAORLFTALPPFEKNCNDNICDDLSITFSFMSLDCLVVG 780	
Db	734	SLVREPIPSQNLRLPVLAVGSQDLFTASLPFEKNCNDNICDDLSITFSFMSLDCLVVG 793	
Qy	781	GRPEENVTVTVNDGEDSVRTQVTFPEPLDLSYRVKSTI-ONORSQSRSLACESASTEV 840	
Db	794	SSLENLVIVVWNAGEDSYGVVSYYPAGLSHRVSGAQKQPHQSALRLACETV-PTED 852	
Qy	841	SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGKLLKANKVTENNMPRTNKEF 900	
Db	853	EG-LASSRCVNHPIFHEGSGTFTVTFDVSYKATLGDRLMLRASASSENKASSKATF 911	
Qy	901	QLELPVKYAVNVVTSHGVSXYLNF-TASENTSRVMOHOYQVSNLQSLPISLVLVP 959	
Db	912	QLELPVKYAVNVVTSHGVSXYLNF-TASENTSRVMOHOYQVSNLQSLPISLVLVP 971	
Qy	960	VLNQTQVWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVYVNCISIAVCORIQCDI 1019	
Db	972	VLLNGVAVVDVMEAPSQSL--PCVSEKPKQHSDFLTQISPSMLDCSIADCLQRCDV 1029	
Qy	1020	PPFGIOREFNATLKNLSPDVIYKTSNHLILVSTAEILPNDISVFTLLPGQGAFFRSQTE 1079	
Db	1030	PSFSVQELDFTLKGNLSFGWVRETLQKKVLVSVSAEITFDTSVYSQLPQGAFFRAQME 1089	
Qy	1080	TKVPEPPEVNPPLTVGVSGVGLLALLITLALYKLGFFKQYKDMMSR 1128	
Db	1090	MVLEDEEVNAPILINGSVGNALLILITLALYKLGFFKQYKDMMSR 1138	
XX	RESULT 14		
XX	AAW23049		
ID	AAW23049	standard; protein; 1161 AA.	
XX	AAW23049;		
XX	24-FEB-1998	(first entry)	
DE	Human beta 2 integrin alpha d subunit.		
KW	Beta 2 integrin alpha d subunit; human; cell migration; cell adhesion; phagocytosis; diabetes; atherosclerosis; multiple sclerosis; asthma; peoriasis; lung inflammation; acute respiratory distress syndrome; rheumatoid arthritis.		
OS	Homo sapiens.		
XX	Key	Location/Qualifiers	
PH	Domain	17..1108	

```

FT      /label= Extracellular_domain
FT      150..352
FT      /notes="region homologous to the I (insertion) domain
FT      common to CD11a, CD11b and CD11c"
FT      Domain
FT      1109..1128
FT      /label= Transmembrane_domain
FT      /notes="homologous to the human CD11c transmembrane
FT      region"
FT      1129..1161
FT      /label= Cytoplasmic_domain
XX      WO9731099-A1.
XX      28-AUG-1997.
XX      24-FEB-1997; 97WO-US002713.
XX      22-FEB-1996; 96US-00605672.
XX      (ICCS-) ICOS CORP.
XX      Gallatin WM, Van Der Vieren M;
XX      WPI; 1997-435154/40.
XX      N-PSDB; AAT79220.
XX      Hybridoma 199M and antibody secreted by it - specific for new rat beta2
XX      integrin subunit, useful to detect subunit in cells and modulate its
XX      activity.
XX      Example 5; Page 116-120; 22pp; English.
XX      This polypeptide comprises a novel human beta 2 integrin subunit,
XX      designated alpha d. Its sequence was deduced from a cDNA clone (see
XX      AAT79220) isolated from a spleen cDNA library. Alpha d is involved in
XX      cell migration, phagocytosis and cell-cell interaction. Recombinant alpha
XX      d polypeptides can be expressed in transformed host cells for use in
XX      assays for identifying antibodies or other compounds that modulate alpha
XX      d activity or which modulate the interaction between alpha d and a
XX      ligand, for treating or preventing diseases in which macrophages are
XX      implicated. Treatment is applicable to disease states in which alpha d
XX      binding, or localised accumulation of cells which express alpha d, is
XX      implicated such as such as type I diabetes, atherosclerosis, multiple
XX      sclerosis, asthma, psoriasis, lung inflammation, acute respiratory
XX      distress syndrome and rheumatoid arthritis
XX      Sequence 1161 AA;
XX      Query Match
XX      Best Local Similarity 58.9%; Pred. No. 2.2e-271;
XX      Matches 665; Conservative 168; Mismatches 288; Indels 8; Gaps 6;
XX      1 FNLDTENAMTFQENARGFGQSVVQLQGRVWVYGAPOEIVAAQORGLYQDYSTGSCPEI 60
XX      17 FKLVDREPTFQEDAGFGQSVVQFGSLVGVGAPLEVVAAQVGTGLYDCAATGMCQPI 76
XX      61 RLQVPVAVNMSLGLAATTPPOLACGPTVHQCSTENTYVKGLCFLFGNLNQPOPK 120
XX      77 PLHIRPEAVNMSLGLTAASTNGSRLLAGCPTLHRVCGENSYSKSGCLLGSRW-ETIQT 135
XX      121 FPEALRGCEQSDIAFLIDGSGSIIPHDFRAMELVSTIMEQLKSKTLPSLMCYSEEF 180
XX      136 VPDATPECHQEMDIVFLIDGSGSDQDNFNQMGFVQAVMGQFEGTDTFLPAMQISNLL 195
XX      181 RHFTTFKQNNPNRSLKPIQLLGRTHRTATGLKVVYRELNFNTNGARKNAKIFLL 240
XX      196 KIHFTTFQRTSPSQSLVDPIVLKGLTFTATGILTVVTLFHHKNGARKSAKKILIVI 255
XX      241 TDGEKGDPLGVEDVITPELDREVIVYVLFGDAFSEKSRQELVASKPRPDHVFQAN 300
XX      256 TDGQKYKPLEVSDVITPQAEKAGIIRYAGVGHAFQGPATROELATISSAPPQDHVPKV 315
XX      301 NFEALTVQNQLREKIFALTEGTQTGSSSFEHMSQEGFSAATISNGPLLSTVGSDVWAG 360

```

```

Db      316 NFAALGSIQKLOEKIYAVECTQSRASSSQEGFSTALTMDFLAGVGSFWSG 375
Qy      361 GVFLVTSKEKSTFNTMTVDSDMDAYLGAAAIILNRVOSLVLGAPRYOHIGLVAMFR 420
Db      376 GAPLYPNMSTFTFNMSENVDMSYLGSTELAKWGVQNVLVGARYOHTGKAVFT 435
Qy      421 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYIYOTRGQGVSVCP 480
Db      436 QVSRQWRKKAETVGTQIGSYFGASLCSVDVDSGSDTLILIGAPHYIYOTRGQGVSVCP 495
Qy      481 PRGORARWQCDVAVLYGEGQPGWGRFGAALTVDLVNMGDKLTDVAIGAPGEBDNRGAVLYF 540
Db      496 PRGORVQWQCDVAVLRGEQGHGPGWGRFGAALTVDLVNMGDKLTDVAIGAPGEBDNRGAVLYF 555
Qy      541 HGTSGSISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGHLVDLTVGAOHHVLLRSQ 600
Db      556 HGASESGISPSHSQRIASSQLSPRLQYFGQSLSGQDLTMDGHLVDLTVGAOHHVLLRS 615
Qy      601 PVLAVKALMBENPREVARNVFECNDQVVKGEAGEVRVCLHVOKSTRDLREGQIQSVVT 660
Db      616 PVLKVGVAWRFSPVEKAVYRCWEEKPSALEAGDATVCLTIQKSSLDQL--GDIOSSVR 673
Qy      661 YDLALDSGRPHSRVAFNETKSTRRTQVGLGTQCTETLKLQLPNCIBDPVSVILRLNF 720
Db      674 FDLALDPCBLTSRALFNETKNTLTTRKTLGLGHCETLKLKLLPCVDEVDVSPILHLNF 733
Qy      721 SLVGTPLSAFGRNRPVLAEDAORLFTALPPEKNCNDNICODDLISITFSNLSLCLVVG 780
Db      734 SLVREPIPSQNLRPVLAAGSDLTASLPEPKNCQDGLCEGLGVTLSPGLQTLTVG 793
Qy      781 GPRFNVTVTVRNDGEDSYRTQVTFPFLDLSYRKVSTLQNRQSRWSRLACESASTEV 840
Db      794 SSLELNVIYTVNAGEDSYGTIVSVLYPAGLSHRRVSGAQKPHQSALRLACETV-PTED 852
Qy      841 SGALKSTSCSNHPIPPENSEVTNITFDVDSKASLGNKLLKANVTSENMPNKTKEF 900
Db      853 EG-LRSSCSNVHPFIPHEGSGTIVTFDVSYKATGLDRMLRMLASSENKASSSKAFF 911
Qy      901 QLELPKVAVTVVTVSHGVSTKYLNLF-TASNTSRVMOHVOYVSNLQGRSIPISLVFLVP 959
Db      912 QLELPKVAVTVVTVSHGVSTKYLNLF-TASNTSRVMOHVOYVSNLQGRSIPISLVFLVP 971
Qy      960 VRLNQTVIWDPRQVTFSENLSSTCTHKEKRLPSHSDFLAELKAPVNVGSIACVQRIQCDI 1019
Db      972 VLLAGVAVWVDMVMEAPSQSL--PCVSEKPPQHSDFLTQISRSPMLDCSIADCLQFCDV 1029
Qy      1020 PPFQIQEZFENATLKNLSFDWYIKTSHNHLIVSTABILFNDSVFTLPLGQGFVRSQTE 1079
Db      1030 PFSVQSELDFTLKNLSFGWVRETLQKKVLVSVVAEITFTSVYSQLPQGEAFWQAWE 1089
Qy      1080 TKVPEFPEYVNPFLPIVSGSVGLLILALITAAALYKLGFFKQYKDMSE 1128
Db      1090 MVLEEDVYNAIPIIMGSSVGALLILALITATLYKLGFFKHYKEMLED 1138
XX      RESULT 15
XX      AAWS7491
XX      ID AAWE7491 standard; protein; 1161 AA.
XX      AC AAWE7491;
XX      DT 24-AUG-1998 (first entry)
XX      DE Human beta2 integrin alpha subunit (alpha d) polypeptide.
XX      KW Beta 2 integrin alpha subunit; alpha d; human; treatment; diabetes;
XX      KW reporter-transactivator construct; atherosclerosis; atherosclerosis;
XX      KW inflammatory bowel disease; arthritis; multiple sclerosis.
XX      OS Homo sapiens.
XX      PH Key Location/Qualifiers

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 17:07:20 ; Search time 18.3484 Seconds
(without alignments)
3199.127 Million cell updates/sec

Title: US-09-902-481b-3

Perfect score: 5879

Sequence: 1 FNLDTENAMTFOENARGFGQ.....FKRQYKDMNSEGPPGAEPO 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/prodata/2/iaa/5A COMB.pcp.*
2: /cgn2_6/prodata/2/iaa/5B COMB.pcp.*
3: /cgn2_6/prodata/2/iaa/6A COMB.pcp.*
4: /cgn2_6/prodata/2/iaa/6B COMB.pcp.*
5: /cgn2_6/prodata/2/iaa/PCTUS COMB.pcp.*
6: /cgn2_6/prodata/2/iaa/backfiles1.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5839	99.3	1153	1	US-08-173-497-3
2	5839	99.3	1153	1	US-08-286-889-3
3	5839	99.3	1153	1	US-08-485-618-3
4	5839	99.3	1153	1	US-08-362-652-3
5	5839	99.3	1153	2	US-08-605-672-3
6	5839	99.3	1153	2	US-08-482-293A-3
7	5839	99.3	1153	2	US-08-943-363-3
8	5839	99.3	1153	3	US-09-193-043-3
9	5839	99.3	1153	4	US-09-688-307A-3
10	5839	99.3	1153	4	US-09-350-259-3
11	5808.5	98.8	1152	2	US-08-476-062A-43
12	5808.5	98.8	1152	5	PCT-US96-01314-43
13	5808.5	98.8	1152	2	5423399-2
14	3446	58.6	1163	2	US-08-476-062A-44
15	3446	58.6	1163	5	PCT-US96-01314-44
16	3423	58.2	1163	1	US-08-173-497-4
17	3423	58.2	1163	1	US-08-286-889-4
18	3423	58.2	1163	1	US-08-485-618-4
19	3423	58.2	1163	1	US-08-362-652-4
20	3423	58.2	1163	2	US-08-605-672-4
21	3423	58.2	1163	2	US-08-482-293A-4
22	3423	58.2	1163	2	US-08-943-363-4
23	3423	58.2	1163	3	US-09-193-043-4
24	3423	58.2	1163	4	US-09-688-307A-4
25	3423	58.2	1163	4	US-09-350-259-4
26	3388	57.6	1161	1	US-08-173-497-2
27	3388	57.6	1161	1	US-08-286-889-2

RESULT 1
US-08-173-497-3
; Sequence 3, Application US/08173497
; Patent No. 5437958
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; TITLE OF INVENTION: No. 5437958el Human 2 Integrin Alpha
; TITLE OF INVENTION: Subunit
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 S. Wacker Drive, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/173,497
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5437958and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 278666/31363
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-173-497-3

ALIGNMENTS

28 3388 57.6 1161 1 US-08-485-618-2 Sequence 2, Appli
29 3388 57.6 1161 1 US-08-362-652-2 Sequence 2, Appli
30 3388 57.6 1161 2 US-08-605-672-2 Sequence 2, Appli
31 3388 57.6 1161 2 US-08-482-293A-2 Sequence 2, Appli
32 3388 57.6 1161 2 US-08-943-363-2 Sequence 2, Appli
33 3388 57.6 1161 3 US-09-193-043-2 Sequence 2, Appli
34 3388 57.6 1161 4 US-09-688-307A-2 Sequence 2, Appli
35 3388 57.6 1161 1 US-09-350-259-2 Sequence 2, Appli
36 3372.5 57.4 1161 1 US-08-485-618-99 Sequence 99, Appli
37 3372.5 57.4 1161 2 US-08-605-672-99 Sequence 99, Appli
38 3372.5 57.4 1161 2 US-08-482-293A-99 Sequence 99, Appli
39 3372.5 57.4 1161 2 US-08-943-363-99 Sequence 99, Appli
40 3372.5 57.4 1161 3 US-09-193-043-99 Sequence 99, Appli
41 3372.5 57.4 1161 4 US-09-688-307A-99 Sequence 99, Appli
42 3372.5 57.4 1161 4 US-09-350-259-99 Sequence 99, Appli
43 3213.5 54.7 1161 3 US-09-193-043-55 Sequence 55, Appli
44 3213.5 54.7 1161 4 US-09-688-307A-55 Sequence 55, Appli
45 3213.5 54.7 1161 4 US-09-350-259-55 Sequence 55, Appli

Query Match 99.3%; Score 5839; DB 1; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVVGAPOEIVAAVNRQSLVQCDYSTGSCBPI 60
DB 17 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVVGAPOEIVAAVNRQSLVQCDYSTGSCBPI 76

QY 61 RLQVPVAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVKGICFLFGSNLRQPOK 120
DB 77 RLQVPVAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVKGICFLFGSNLRQPOK 136
QY 121 FPEARLGCPOEDSDIAFLIDGSGSIIPHDPRMKELVSTIMEOLKSKTFLSLMOYSEEF 180
DB 137 FPEARLGCPOEDSDIAFLIDGSGSIIPHDPRMKEFVSTVMEOLKSKTFLSLMOYSEEF 196
QY 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVVRRELNTNGARKNAFKILFLL 240
DB 197 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVVRRELNTNGARKNAFKILVVI 256
QY 241 TDGEKFGDPLGYEDVPEADREGVIRVYVGLGFAFRSEKSRQELNTVASPPRHHVFOAN 300
DB 257 TDGEKFGDPLGYEDVPEADREGVIRVYVGLGFAFRSEKSRQELNTIASKPPRHHVFOAN 316
QY 301 NFALKTQVQLREKIPALISGTQTGSSSPHEHMSQEGFSAATISNGPLLSSTVGSYDWAG 360
DB 317 NFALKTQVQLREKIPALISGTQTGSSSPHEHMSQEGFSAATISNGPLLSSTVGSYDWAG 376
QY 361 GVFLYTSKSKSTFINMTRVDSMDNDAYLGYAAALILNRVQSLVLAGAPRYOHIGLVAMER 420
DB 377 GVFLYTSKSKSTFINMTRVDSMDNDAYLGYAAALILNRVQSLVLAGAPRYOHIGLVAMER 436
QY 421 QNTGMWESNANVKGTOIGAFGASLCSVDVDSNGSTDVLVIGAPHYYEQTRGGQVSVCPIL 480
DB 437 QNTGMWESNANVKGTOIGAFGASLCSVDVDSNGSTDVLVIGAPHYYEQTRGGQVSVCPIL 496
QY 481 PRGORARWQCDVLYGEGOPWKGFGAALTVLGDVNGDKLTVAI GARGEDNRGAVILF 540
DB 497 PRGORARWQCDVLYGEGOPWKGFGAALTVLGDVNGDKLTVAI GARGEDNRGAVILF 556
QY 541 HGTSGSGISPSHSORJAGSKSLPRLOVFGQSLGGQDLTMDGLVDTLVGAQGHVLLRSQ 600
DB 557 HGTSGSGISPSHSORJAGSKSLPRLOVFGQSLGGQDLTMDGLVDTLVGAQGHVLLRSQ 616
QY 601 PVLRLVXALMFNPREVARNVFCNDQVYKGEAGEVRVCLHVOKSTRDLREQIQSVVT 660
DB 617 PVLRLVXALMFNPREVARNVFCNDQVYKGEAGEVRVCLHVOKSTRDLREQIQSVVT 676
QY 661 YDLALDSGRPHSAVFNETKSNTRRQTVLGLTQTCETLKLQPNICIEDPVPSPVILRLNF 720
DB 677 YDLALDSGRPHSAVFNETKSNTRRQTVLGLTQTCETLKLQPNICIEDPVPSPVILRLNF 736
QY 721 SLVGTPLSAPGNLRPVLAEDAQLFTALPPFEXKNCNDNICQDDLSITFSPMSLDCLVWG 780
DB 737 SLVGTPLSAPGNLRPVLAEDAQLFTALPPFEXKNCNDNICQDDLSITFSPMSLDCLVWG 796
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSQRSLACESASSTEV 840
DB 797 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSQRSLACESASSTEV 856
QY 841 SGALKSTSCSINHPIIPENSEVTFNITFDVDSKASLGKLLKANTYSENMPRTNKTFF 900
DB 857 SGALKSTSCSINHPIIPENSEVTFNITFDVDSKASLGKLLKANTYSENMPRTNKTFF 916
QY 901 QLELPVKYAVYVMTSHGVSTKYLNPTASNTSRVMOHQYVSNLQORSIPISLVLPLVPV 960
DB 917 QLELPVKYAVYVMTSHGVSTKYLNPTASNTSRVMOHQYVSNLQORSIPISLVLPLVPV 976
QY 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCQRIQCDDIP 1020
DB 977 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCQRIQCDDIP 1036
QY 1021 FFGIQEFNATLKNLSFDWYIKTSENHLLIVSTABILNDSVFTLLPCQGAFRVSQTF 1080
DB 1037 FFGIQEFNATLKNLSFDWYIKTSENHLLIVSTABILNDSVFTLLPCQGAFRVSQTF 1096
QY 1081 KVEPFFVPNPLPLIVGSSVGGILLALLIITAAALYKLGFFRQVQKDMWSEGGPPGABPQ 1137
DB 1097 KVEPFFVPNPLPLIVGSSVGGILLALLIITAAALYKLGFFRQVQKDMWSEGGPPGABPQ 1153

RESULT 2

US-08-286-889-3
; Sequence 3, Application US/08286889
; Patent No. 5470953
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Mich
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5470953el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/286,889
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: P38,659
; REFERENCE/DOCKET NUMBER: 27866/32168
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-286-889-3

Query Match 99.3%; Score 5839; DB 1; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 ENLDTENAMTFOENARGFGQSVVQLQGSRVVVGAPQBIVAANORGSLYCCDYSTGSCCEPI 60
DB 17 ENLDTENAMTFOENARGFGQSVVQLQGSRVVVGAPQBIVAANORGSLYCCDYSTGSCCEPI 76
QY 61 RLQVPVAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVKGICFLFGSNLRQPOK 120
DB 77 RLQVPVAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVKGICFLFGSNLRQPOK 136
QY 121 FPEARLGCPOEDSDIAFLIDGSGSIIPHDPRMKELVSTIMEOLKSKTFLSLMOYSEEF 180
DB 137 FPEARLGCPOEDSDIAFLIDGSGSIIPHDPRMKEFVSTVMEOLKSKTFLSLMOYSEEF 196
QY 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVVRRELNTNGARKNAFKILFLL 240
DB 197 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVVRRELNTNGARKNAFKILVVI 256
QY 241 TDGEKFGDPLGYEDVPEADREGVIRVYVGLGFAFRSEKSRQELNTVASPPRHHVFOAN 300
DB 257 TDGEKFGDPLGYEDVPEADREGVIRVYVGLGFAFRSEKSRQELNTIASKPPRHHVFOAN 316
QY 301 NFALKTQVQLREKIPALISGTQTGSSSPHEHMSQEGFSAATISNGPLLSSTVGSYDWAG 360
DB 317 NFALKTQVQLREKIPALISGTQTGSSSPHEHMSQEGFSAATISNGPLLSSTVGSYDWAG 376

QY 361 GVFLYTSKEKSTPINNTRVDSMDNDAYLGAAIILRNVRQSLVIGAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTPINNTRVDSMDNDAYLGAAIILRNVRQSLVIGAPRYQHIGLVAMFR 436
QY 421 QNTGWSNANVKGQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTGRGGQSVQVPL 480
Db 437 QNTGWSNANVKGQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTGRGGQSVQVPL 496
QY 481 PRGORARWQCDVLYGEGQGPNGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db 497 PRGORARWQCDVLYGEGQGPNGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556
QY 541 HTSGSGISPSHSQRIAGSKSPRIQYFQSGSLGGQDLTMDGLVLTGAGQGHVLLRSQ 600
Db 557 HTSGSGISPSHSQRIAGSKSPRIQYFQSGSLGGQDLTMDGLVLTGAGQGHVLLRSQ 616
QY 601 PVLRYKAIMFNPRVARNVFCNDQVVKGEAGVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLRYKAIMFNPREVARNVFCNDQVVKGEAGVRVCLHVQKSTRDLREGQIQSVVT 676
QY 661 YDLALDSGRPHSRVFNKSTRTQOTVLGTQTCETLKLQLPNCIEDPVSPIVLRNF 720
Db 677 YDLALDSGRPHSRVFNKSTRTQOTVLGTQTCETLKLQLPNCIEDPVSPIVLRNF 736
QY 721 SLVGTPLSAGNLRPVLAEDAQRLLFTALPPFEKNCNDNI CDDLSITFSFMSLDCLVYG 780
Db 737 SLVGTPLSAGNLRPVLAEDAQRLLFTALPPFEKNCNDNI CDDLSITFSFMSLDCLVYG 796
QY 781 GRPFRNVTVVRNDCGDSVRQVTFEPLDLSYRKVSTLQNRORSWRLACESASSTEV 840
Db 797 GRPFRNVTVVRNDCGDSVRQVTFEPLDLSYRKVSTLQNRORSWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPKNTKEP 900
Db 857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPKNTKEP 916
QY 901 QLELPVKAVMVTSHGVSTKYLNFTASNTSRVMQHOYQVSNLQORSLPISLVLVPV 960
Db 917 QLELPVKAVMVTSHGVSTKYLNFTASNTSRVMQHOYQVSNLQORSLPISLVLVPV 976
QY 961 RLNQTVIWDPRQVTFSENLSSCTCKERLPKSHDPLAELRKAPVVCNSIACVQRIQCDIP 1020
Db 977 RLNQTVIWDPRQVTFSENLSSCTCKERLPKSHDPLAELRKAPVVCNSIACVQRIQCDIP 1036
QY 1021 PFGIOEENFNLKGNLSFDWIKTSHNHLIVSTHAILFNDSVFTLLPGQAFVRSQTEP 1080
Db 1037 PFGIOEENFNLKGNLSFDWIKTSHNHLIVSTHAILFNDSVFTLLPGQAFVRSQTEP 1096
QY 1081 KVEPPEVNPPLIVGSSVGGLLLALITAAALYKLGPFKROYKDMMSGGPPGABPO 1137
Db 1097 KVEPPEVNPPLIVGSSVGGLLLALITAAALYKLGPFKROYKDMMSGGPPGABPO 1153

RESULT 3
US-08-485-618-3
; Sequence 3, Application US/08485618
; Patent No. 5728533
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5728533el Humar 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485.618
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32797
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1153 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-485-618-3

Query Match 99.3%; Score 5839; DB 1; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;
QY 1 FNLDTENAMTFQENARGFGQSVVQLQSGRVVVGAPQBIIVAAQORGSILYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFQENARGFGQSVVQLQSGRVVVGAPQBIIVAAQORGSILYQCDYSTGSCPEI 76
QY 61 RLQVPVEAVNMSGLSLAATTPPQLLACGPTVHQTCSNTYVVKELCFLFGSNLRQOPQK 120
Db 77 RLQVPVEAVNMSGLSLAATTPPQLLACGPTVHQTCSNTYVVKELCFLFGSNLRQOPQK 136
QY 121 FPEALRCQPEDSDIAPLIDGSGSIIPHDPRMKELVSTIMBOLKSKTFLSLMOYSEEP 180
Db 137 FPEALRCQPEDSDIAPLIDGSGSIIPHDPRMKELVSTIMBOLKSKTFLSLMOYSEEP 196
QY 181 RIHFTFKFQNNPNRSLIKPIITQLLGRTHATGLRKVVRELFNITNGARKNAFKILFL 240
Db 197 RIHFTFKFQNNPNRSLIKPIITQLLGRTHATGLRKVVRELFNITNGARKNAFKILVVI 256
QY 241 TDGEKFGDPLGYEDVTPELDRSGVIRYVLGFGDAPRSEKSRQBELNTVASPPDRHVPQAN 300
Db 257 TDGEKFGDPLGYEDVTPEADREGVIRYVIGVGDAFRSEKSRQBELNTIASKPPDRHVPQVN 316
QY 301 NFEALKTQNLREKIPAIERTGTGSSSPHEHMSQEGPSAALTNSGPLLSTVGSYDWAG 360
Db 317 NFEALKTQNLREKIPAIERTGTGSSSPHEHMSQEGPSAALTNSGPLLSTVGSYDWAG 376
QY 361 GVFLYTSKEKSTPINNTRVDSMDNDAYLGAAIILRNVRQSLVIGAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTPINNTRVDSMDNDAYLGAAIILRNVRQSLVIGAPRYQHIGLVAMFR 436
QY 421 QNTGWSNANVKGQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTGRGGQSVQVPL 480
Db 437 QNTGWSNANVKGQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTGRGGQSVQVPL 496
QY 481 PRGORARWQCDVLYGEGQGPNGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db 497 PRGORARWQCDVLYGEGQGPNGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556
QY 541 HTSGSGISPSHSQRIAGSKSPRIQYFQSGSLGGQDLTMDGLVLTGAGQGHVLLRSQ 600

Db 557 HGTSGGSGISHSQRTAGSKLSPLOYFQGLSGQDLTWDGLVDTVGAQHVLRLRQ 616
Qy 601 PVLVRKAIEMFNPREVARNVFCNDQVYKGEAGVRVCLHVQKSTRDLRBCQIQSVVT 660
Db 617 PVLVRKAIEMFNPREVARNVFCNDQVYKGEAGVRVCLHVQKSTRDLRBCQIQSVVT 676
Qy 661 YDLALDSGRPHSRVAVNETKNSRRTQVGLTQTCETLKLQLPNCIBDPVSPVILRLNP 720
Db 677 YDLALDSGRPHSRVAVNETKNSRRTQVGLTQTCETLKLQLPNCIBDPVSPVILRLNP 736
Qy 721 SLVGTPLSAFNGLRPVLAEADAQLFTALPPFERKNCNDNICODDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFNGLRPVLAEADAQLFTALPPFERKNCNDNICODDLSITFSFMSLDCLVVG 796
Qy 781 GPREFNVTWVRNDGDSYRTQVTPPEPDLISYRKUSTLQKRSQSWLACESASTEV 840
Db 797 GPREFNVTWVRNDGDSYRTQVTPPEPDLISYRKUSTLQKRSQSWLACESASTEV 856
Qy 841 SGALKSTCSINHPPIPNSEVFNITFDVDSKASLGNKLLKANVTSENNPRTNKTEF 900
Db 857 SGALKSTCSINHPPIPNSEVFNITFDVDSKASLGNKLLKANVTSENNPRTNKTEF 916
Qy 901 QLELPKYAYVMVTSHGVSUKYINFTASNTSRVQHQVQVSNLCQSLPISLVLVVP 960
Db 917 QLELPKYAYVMVTSHGVSUKYINFTASNTSRVQHQVQVSNLCQSLPISLVLVVP 976
Qy 961 RLQNTWIDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCSTAVCQRIQCDIP 1020
Db 977 RLQNTWIDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCSTAVCQRIQCDIP 1036
Qy 1021 FPGIOBEFNATLKNLSFDWYIYKTSNNHLLIVSTABILFNDSTFTLLPQGGAFVRSQTF 1080
Db 1037 FPGIOBEFNATLKNLSFDWYIYKTSNNHLLIVSTABILFNDSTFTLLPQGGAFVRSQTF 1096
Qy 1081 KVEPFEVNPPLIVSSVGLLLALITAAALYKLGPFKQYKDMSEGGPPGABPQ 1137
Db 1097 KVEPFEVNPPLIVSSVGLLLALITAAALYKLGPFKQYKDMSEGGPPGABPQ 1153

RESULT 4
US-08-362-652-3
Sequence 3, Application US/08362652
Patent No. 5766850
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. 5766850el Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSES: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 5300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/362,652
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.

REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32391
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1153 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-362-652-3
Query Match 99.3%; Score 5839; DB 1; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;
Qy 1 FNLDTENAMTQENARGFGQSVQVQGSRRVVVVGAPQETVAANORGSLYQCDYSTGSCSCEPI 60
Db 17 FNLDTENAMTQENARGFGQSVQVQGSRRVVVVGAPQETVAANORGSLYQCDYSTGSCSCEPI 76
Qy 61 PLOYPVEAVNMSLGLSLAATTSPOLLACGPTVHTCSENTYVVKLCFLPGSNLRQOQOK 120
Db 77 RLQVPEVAVNMSLGLSLAATTSPOLLACGPTVHTCSENTYVVKLCFLPGSNLRQOQOK 136
Qy 121 FPEALRGCPQSDSDIAFLIDGSGSIIPHDFFRMKELVSTIMEQLKSKTFLSLMOYSSEF 180
Db 137 FPEALRGCPQSDSDIAFLIDGSGSIIPHDFFRMKELVSTIMEQLKSKTFLSLMOYSSEF 196
Qy 181 RIHFTFKFQNNPNSRLIKDITOLLGHTHTATGKRVKRVRELFNITNGARKNAKFLVLL 240
Db 197 RIHFTFKFQNNPNSRLIKDITOLLGHTHTATGKRVKRVRELFNITNGARKNAKFLVLL 256
Qy 241 TDGEKFGPLGVEDVPEADREGVIRYVIGVGDAPFRSEKSRQELNVTASKPRDRHVFOVN 300
Db 257 TDGEKFGPLGVEDVPEADREGVIRYVIGVGDAPFRSEKSRQELNVTASKPRDRHVFOVN 316
Qy 301 NPEALKTVONLRKFIPIAETGTGSSSSFEHMSQEGSAITNSNGPLLSITVGSYDWAG 360
Db 317 NPEALKTVONLRKFIPIAETGTGSSSSFEHMSQEGSAITNSNGPLLSITVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAILNRNRVQSLVLCAPRYOHIGLVAMPR 420
Db 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAILNRNRVQSLVLCAPRYOHIGLVAMPR 436
Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLICAPHYVOTRGGQVSVCP 480
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLICAPHYVOTRGGQVSVCP 496
Qy 481 PRGQARWQCDVAVLYGEOGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db 497 PRGQARWQCDVAVLYGEOGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556
Qy 541 HGTSGGSGISPSHSQRTAGSKLSPLOYFQGLSGQDLTMDGLVDTVGAQHVLRLRQ 600
Db 557 HGTSGGSGISPSHSQRTAGSKLSPLOYFQGLSGQDLTMDGLVDTVGAQHVLRLRQ 616
Qy 601 PVLVRKAIEMFNPREVARNVFCNDQVYKGEAGVRVCLHVQKSTRDLRBCQIQSVVT 660
Db 617 PVLVRKAIEMFNPREVARNVFCNDQVYKGEAGVRVCLHVQKSTRDLRBCQIQSVVT 676
Qy 661 YDLALDSGRPHSRVAVNETKNSRRTQVGLTQTCETLKLQLPNCIBDPVSPVILRLNP 720
Db 677 YDLALDSGRPHSRVAVNETKNSRRTQVGLTQTCETLKLQLPNCIBDPVSPVILRLNP 736
Qy 721 SLVGTPLSAFNGLRPVLAEADAQLFTALPPFERKNCNDNICODDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFNGLRPVLAEADAQLFTALPPFERKNCNDNICODDLSITFSFMSLDCLVVG 796
Qy 781 GPREFNVTWVRNDGDSYRTQVTPPEPDLISYRKUSTLQKRSQSWLACESASTEV 840

Db 797 GPREFNVTVVRNDGEDSVRTQVTFPPFLDLSYRKVSTLQNRQSRWELACESASSTEV 856
Qy 841 SGALKSTCSINHPISPESEVFNITFDVDSKASIGNKLLKANKVTSNNMPTNKTEF 900
Db 857 SGALKSTCSINHPISPESEVFNITFDVDSKASIGNKLLKANKVTSNNMPTNKTEF 916
Qy 901 QLELPVKYAVYVMTSHGVSTKYNLTASNTSRVMQHOVQVSNLQORSIPISLVLFPV 960
Db 917 QLELPVKYAVYVMTSHGVSTKYNLTASNTSRVMQHOVQVSNLQORSIPISLVLFPV 976
Qy 961 RLNQTVINDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCSTAVCQRIQCDIP 1020
Db 977 RLNQTVINDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCSTAVCQRIQCDIP 1036
Qy 1021 FPGIOSEFNATKGNLSFDWYIKTSHNLLIVSTAEILFNDVSPTLLPGQGFVRSOTET 1080
Db 1037 FPGIOSEFNATKGNLSFDWYIKTSHNLLIVSTAEILFNDVSPTLLPGQGFVRSOTET 1096
Qy 1081 KVEPPEVNPPLPLIVSSVGLLLALITAAALYKLGFFKQYKDMKSEGGPPGABEQ 1137
Db 1097 KVEPPEVNPPLPLIVSSVGLLLALITAAALYKLGFFKQYKDMKSEGGPPGABEQ 1153

RESULT 5

US-08-605-672-3
; Sequence 3, Application US/08605672
; Patent No. 5817515
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5817515el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/605,672
; FILING DATE:
; CLASSIFICATION: 530
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32684
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein

US-08-605-672-3

Query Match 99.3%; Score 5839; DB 2; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;
Qy 1 FNLDTENAMTFOENARGFGQSVVQLQGSRVVVGAPQEIIVAAQNRQSLVQCDYSTGSCBPI 60
Db 17 FNLDTENAMTFOENARGFGQSVVQLQGSRVVVGAPQEIIVAAQNRQSLVQCDYSTGSCBPI 76
Qy 61 RLOVPEAVNNLSGLSLAANTSPOLLACGPTVHOTCSNTVYKGLCFGLFNSLQOQPK 120
Db 77 RLOVPEAVNNLSGLSLAANTSPOLLACGPTVHOTCSNTVYKGLCFGLFNSLQOQPK 136
Qy 121 FPEALRGCPQEDSDIAFLIDGSGSIIPDPRMKELVSTIMEQLKSKTLPMSLMSYSEF 180
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIPDPRMKELVSTIMEQLKSKTLPMSLMSYSEF 196
Qy 181 RIHPTFKFQNNPNSLIKGITOLGRTHTATGKRVVRELFNITNGARKNAKILPL 240
Db 197 RIHPTFKFQNNPNSLIKGITOLGRTHTATGKRVVRELFNITNGARKNAKILPL 256
Qy 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAFRSEKSRQELNTVASKPPRDHVFQAN 300
Db 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAFRSEKSRQELNTVASKPPRDHVFQAN 316
Qy 301 NFEALKTQVQLREKIPAIETGTCSSSSFEHMSQSGSAITNGPLLSVTGSDWAG 360
Db 317 NFEALKTQVQLREKIPAIETGTCSSSSFEHMSQSGSAITNGPLLSVTGSDWAG 376
Qy 361 GYPLVTSSEKSTFINMTRVDSMDNDAYLGYAAAIILNRVQSLVLCAPRYQHIGLVAMFR 420
Db 377 GYPLVTSSEKSTFINMTRVDSMDNDAYLGYAAAIILNRVQSLVLCAPRYQHIGLVAMFR 436
Qy 421 QNTGWSNNANVKGTOIGAYFGASLCSVDVDSNGSTDVLICAPHYETRGQGVSVCP 480
Db 437 QNTGWSNNANVKGTOIGAYFGASLCSVDVDSNGSTDVLICAPHYETRGQGVSVCP 496
Qy 481 PRGQARWQCDVAVLGEQGPWRFGAALTVLGDVNGDKLTDVAIGAPCEENRGAVILF 540
Db 497 PRGQARWQCDVAVLGEQGPWRFGAALTVLGDVNGDKLTDVAIGAPCEENRGAVILF 556
Qy 541 HGTSGSISPSHSQRIAGSKLSPRIQYFGQSLSGQDLTMDGLVLTGCAQHVLLLRQ 600
Db 557 HGTSGSISPSHSQRIAGSKLSPRIQYFGQSLSGQDLTMDGLVLTGCAQHVLLLRQ 616
Qy 601 PVLRVKATMEFNPREVARNVPECNDQVVKGEAGVRVCLHVQKSTRDLRSGQIQSVVT 660
Db 617 PVLRVKATMEFNPREVARNVPECNDQVVKGEAGVRVCLHVQKSTRDLRSGQIQSVVT 676
Qy 661 YDLALDSGRPHSRAVENETKSTRQTVLGLTQTCETLKLQLPNCIEDPVSPVILRLNF 720
Db 677 YDLALDSGRPHSRAVENETKSTRQTVLGLTQTCETLKLQLPNCIEDPVSPVILRLNF 736
Qy 721 SLVGTPLSAFGNLRPVLAEDAQRLLFTALFPPEKNGNDNICODDLSITFESMSLCLVVG 780
Db 737 SLVGTPLSAFGNLRPVLAEDAQRLLFTALFPPEKNGNDNICODDLSITFESMSLCLVVG 796
Qy 781 GPREFNVTVVRNDGEDSVRTQVTFPPFLDLSYRKVSTLQNRQSRWELACESASSTEV 840
Db 797 GPREFNVTVVRNDGEDSVRTQVTFPPFLDLSYRKVSTLQNRQSRWELACESASSTEV 856
Qy 841 SGALKSTCSINHPISPESEVFNITFDVDSKASIGNKLLKANKVTSNNMPTNKTEF 900
Db 857 SGALKSTCSINHPISPESEVFNITFDVDSKASIGNKLLKANKVTSNNMPTNKTEF 916
Qy 901 QLELPVKYAVYVMTSHGVSTKYNLTASNTSRVMQHOVQVSNLQORSIPISLVLFPV 960
Db 917 QLELPVKYAVYVMTSHGVSTKYNLTASNTSRVMQHOVQVSNLQORSIPISLVLFPV 976
Qy 961 RLNQTVINDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCSTAVCQRIQCDIP 1020
Db 977 RLNQTVINDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCSTAVCQRIQCDIP 1036

QY 1021 PFGIOEBFNATLKGNSLSPDVIKTSNHLILVSTABILFNDVSFTLLPGQAFVRSQTET 1080
DB 1037 PFGIOEBFNATLKGNSLSPDVIKTSNHLILVSTABILFNDVSFTLLPGQAFVRSQTET 1096
QY 1081 KVBEPFVNPPLIVGSSVGLLALITAAALYKLGFFKQYKQKDMSEGGPPGAEPO 1137
DB 1097 KVBEPFVNPPLIVGSSVGLLALITAAALYKLGFFKQYKQKDMSEGGPPGAEPO 1153

RESULT 6

US-08-482-293A-3
; Sequence 3, Application US/08482293A
; Patent No. 5831029
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5831029el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,293A
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 278666/32584
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein

US-08-482-293A-3

Query Match 99.3%; Score 5839; DB 2; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 ENLDTENAMTFENARGFQSVVQLQGRVVGAPQEIIVANORGLYQCDYSTGSCPEI 60
DB 17 ENLDTENAMTFENARGFQSVVQLQGRVVGAPQEIIVANORGLYQCDYSTGSCPEI 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSENTYVKGKLCFLFGSNLRQQPK 120
DB 77 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSENTYVKGKLCFLFGSNLRQQPK 136
QY 121 PFALRGCPQEDSDIAFLIDGSGSIIPHDFRMKBEFVSTVMEQLKSKTFLSLMQYSEEP 180

RESULT 7

US-08-943-363-3
; Sequence 3, Application US/08943363
; Patent No. 5837478
; GENERAL INFORMATION:

APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. 5837478e1 Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSES: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICANT: US/08/943,363
FILING DATE: 5-AUG-1994
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,839
FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32684
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1153 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-943-363-3

Query Match
Best local Similarity 99.3%; Score 5839; DB 2; Length 1153;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQEIIVAAQSGSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQEIIVAAQSGSLYQCDYSTGSCPEI 76

QY 61 RLQVPVAVNMSGLSLAATTSPPQLACGPTVHQCSTENTYKGLCFPLGSLNRQOPK 120
DB 77 RLQVPVAVNMSGLSLAATTSPPQLACGPTVHQCSTENTYKGLCFPLGSLNRQOPK 136

QY 121 FPEALRCPOBDSIAFLIDGSGIIPHDFFRMKELVSTIMEOLKSKTILFSLMYSSEEP 180
DB 137 FPEALRCPOBDSIAFLIDGSGIIPHDFFRMKELVSTIMEOLKSKTILFSLMYSSEEP 196

QY 181 RIHPTFKFQNNPRLIPIKPIITOLLGRTATGLRKVRVRELFININGARKNAFKILFL 240
DB 197 RIHPTFKFQNNPRLIPIKPIITOLLGRTATGLRKVRVRELFININGARKNAFKILVI 256

QY 241 TCEKFGDPLGYEDVIEPDLREGVIRVVLGFGDAFSEKSRQELNTVASPPRDHVFQAN 300
DB 257 TDGKFGDPLGYEDVIEPDLREGVIRVVLGFGDAFSEKSRQELNTVASPPRDHVFQAN 316

QY 301 NFEALKTQNLREKIFAIECTQTGSSSSSEHEMSQSGFAAITSNGPLSTVGSYDWAG 360
DB 317 NFEALKTQNLREKIFAIECTQTGSSSSSEHEMSQSGFAAITSNGPLSTVGSYDWAG 376

RESULT 8

US-09-193-043-3
; Sequence 3, Application US/09193043
; Patent No. 6251395
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 6251395e1 Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/193,043
; EARLIER FILING DATE: 1998-11-16
; EARLIER FILING DATE: 1993-12-23
; EARLIER FILING DATE: 1993-12-23
; EARLIER FILING DATE: 1994-08-05
; EARLIER FILING DATE: 1994-08-05
; EARLIER FILING DATE: 1994-12-21
; EARLIER FILING DATE: 1997-10-03

QY 361 GVFLYTSKEKSTFINMTRVSDMNDAYLGAAAIILNRVQSLVLAGAPRYOHIGLVAMFR 420
DB 377 GVFLYTSKEKSTFINMTRVSDMNDAYLGAAAIILNRVQSLVLAGAPRYOHIGLVAMFR 436

QY 421 QNTGWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYBQYQOTGGQSVQCLP 480
DB 437 QNTGWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYBQYQOTGGQSVQCLP 496

QY 481 PRQQRARWQCDVLYGEGQQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
DB 497 PRQQRARWQCDVLYGEGQQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDCLVDLTGAGCHVLLRSQ 600
DB 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDCLVDLTGAGCHVLLRSQ 616

QY 601 PVLRVKAIMFNPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
DB 617 PVLRVKAIMFNPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676

QY 661 YDLALDSGRPHSRVFNETKSTTRQTVLGLTQTCETLKLQENCIEDPSPVILRLNF 720
DB 677 YDLALDSGRPHSRVFNETKSTTRQTVLGLTQTCETLKLQENCIEDPSPVILRLNF 736

QY 721 SLVGTPLSAPGNLAPVLAEDAQRFTALFPPEKKGNDNICODDLSITFSFMSLDCLVVG 780
DB 737 SLVGTPLSAPGNLAPVLAEDAQRFTALFPPEKKGNDNICODDLSITFSFMSLDCLVVG 796

QY 781 GPREFNVTTVVRNDGSDSYRTQVTFPPFLDISYRKVSTLQNRQSRWSRLACESASSTEV 840
DB 797 GPREFNVTTVVRNDGSDSYRTQVTFPPFLDISYRKVSTLQNRQSRWSRLACESASSTEV 856

QY 841 SGALKSTSCSINHPIFPENSESVTFNITFDVDSKSLGKLLKANVTSENNMPTNKTEF 900
DB 857 SGALKSTSCSINHPIFPENSESVTFNITFDVDSKSLGKLLKANVTSENNMPTNKTEF 916

QY 901 QLELPVKYAVVMVTVSHGVSTKYLNTASNTSRVMQHQYQVSNIGQSLPISLVLFPV 960
DB 917 QLELPVKYAVVMVTVSHGVSTKYLNTASNTSRVMQHQYQVSNIGQSLPISLVLFPV 976

QY 961 RLNQTVIWDROVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCSTAVCORIQCDIP 1020
DB 977 RLNQTVIWDROVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCSTAVCORIQCDIP 1036

QY 1021 FFGIOEFNATLKGNLSPDWIKTSHNHLIVSTAEIIFNDSVFTLLPQCGAFVRSQSTET 1080
DB 1037 FFGIOEFNATLKGNLSPDWIKTSHNHLIVSTAEIIFNDSVFTLLPQCGAFVRSQSTET 1096

QY 1081 KVEPFEVNPPLIIVGSSVGGLLILALITAAALYKLGPPKQYKDMMSGGPPGABPQ 1137
DB 1097 KVEPFEVNPPLIIVGSSVGGLLILALITAAALYKLGPPKQYKDMMSGGPPGABPQ 1153

Db 317 NFEALKTIONQREKIFAIEGTQGTSSSFEHMSQSGFSAAITNSGPLLSTVGSYDWAG 376
Qy 361 GYFLYTSKSTFIINWTRVDSMDNDAYLGAAAIILNRVQSLVGLGAPRYOHIGLVAMER 420
Db 377 GYFLYTSKSTFIINWTRVDSMDNDAYLGAAAIILNRVQSLVGLGAPRYOHIGLVAMER 436
Qy 421 QNTGWMESNANVKGTOIGAYFGASLCSDVDVNGSTDLVLI GAPHYETRGQGVSVCP 480
Db 437 QNTGWMESNANVKGTOIGAYFGASLCSDVDVNGSTDLVLI GAPHYETRGQGVSVCP 496
Qy 481 PRGQARWQCDVAVLGEQCPWGRFGAALTVDVNGDKLTDVAIGAPCEEDNRGAVYLF 540
Db 497 PRGQARWQCDVAVLGEQCPWGRFGAALTVDVNGDKLTDVAIGAPCEEDNRGAVYLF 556
Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDTVGAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDTVGAQGHVLLRSQ 616
Qy 601 PVLRVKAIEMFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLRRCQIOSVVT 660
Db 617 PVLRVKAIEMFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLRRCQIOSVVT 676
Qy 661 YDLALDSGRPHSRVAFNETKNSTRRTQVLTGTCTETLKLQLPNCIEDPVSPIVLRNF 720
Db 677 YDLALDSGRPHSRVAFNETKNSTRRTQVLTGTCTETLKLQLPNCIEDPVSPIVLRNF 736
Qy 721 SLVGTPLSAFNLRPVLAEDAORLTALPPFEKNCNDNICODDLSITFSFMSLDCLVWG 780
Db 737 SLVGTPLSAFNLRPVLAEDAORLTALPPFEKNCNDNICODDLSITFSFMSLDCLVWG 796
Qy 781 GPREFNVTVVRNDGEDSVYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 840
Db 797 GPREFNVTVVRNDGEDSVYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 856
Qy 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGKLLKANKYTSNNMPTNKTEF 900
Db 857 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGKLLKANKYTSNNMPTNKTEF 916
Qy 901 QLELPVKIAYVWVTSHGVSSTKYINFTASNTSRVMOHQVOVSNLQORSLSPLSLVFLPV 960
Db 917 QLELPVKIAYVWVTSHGVSSTKYINFTASNTSRVMOHQVOVSNLQORSLSPLSLVFLPV 976
Qy 961 RLNQTVIWDSPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCISIAVCORIQCIP 1020
Db 977 RLNQTVIWDSPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCISIAVCORIQCIP 1036
Qy 1021 PFGIOBEFNATLKNLSFDVYIKTSHNHLIVSTAEILFNDVSFTLLPGQGAFFRSQTTET 1080
Db 1037 PFGIOBEFNATLKNLSFDVYIKTSHNHLIVSTAEILFNDVSFTLLPGQGAFFRSQTTET 1096
Qy 1081 KVEPPEFVNPPLITVSGSSVGLLALLALITAAALKLGFPRQYKDMMSBEGPPGABPQ 1137
Db 1097 KVEPPEFVNPPLITVSGSSVGLLALLALITAAALKLGFPRQYKDMMSBEGPPGABPQ 1153

RESULT 10
US-09-350-259-3
; Sequence 3, Application US/09350259
; Patent No. 6620915
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; TITLE OF INVENTION: No. 6620915el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350,259
; CURRENT FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: 09/193,043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652

; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-259-3

Query Match 99.3%; Score 5839; DB 4; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

Qy 1 FNLDPTENAMTQENARGFGQSVVQLQGSRVVVGAPOEIIVAAQNRGSLVQCDYSTSCBPI 60
Db 17 FNLDPTENAMTQENARGFGQSVVQLQGSRVVVGAPOEIIVAAQNRGSLVQCDYSTSCBPI 76
Qy 61 RLOVPEAVNMSLGLSLAATTSPOLLACGPTVHTCSENTVVKGLCFLPGSNLRFQOQPK 120
Db 77 RLOVPEAVNMSLGLSLAATTSPOLLACGPTVHTCSENTVVKGLCFLPGSNLRFQOQPK 136
Qy 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPFRMKELVSTIMEQLKSKTILFSLMOYSBEP 180
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPFRMKELVSTIMEQLKSKTILFSLMOYSBEP 196
Qy 181 RIHPTFKFQONPNPERSLIPITQLLGRTHATGLRKVVRELFNITNGARKNAKILPVL 240
Db 197 RIHPTFKFQONPNPERSLIPITQLLGRTHATGLRKVVRELFNITNGARKNAKILPVL 256
Qy 241 TDGERFGPGLGYEDVPELDRGVIRYVILGFDGAFSEKSRQELNVTASKPPOHVFQAN 300
Db 257 TDGERFGPGLGYEDVPELDRGVIRYVILGFDGAFSEKSRQELNVTASKPPOHVFQAN 316
Qy 301 NFEALKTIONQREKIFAIEGTQGTSSSFEHMSQSGFSAAITNSGPLLSTVGSYDWAG 360
Db 317 NFEALKTIONQREKIFAIEGTQGTSSSFEHMSQSGFSAAITNSGPLLSTVGSYDWAG 376
Qy 361 GYFLYTSKSTFIINWTRVDSMDNDAYLGAAAIILNRVQSLVGLGAPRYOHIGLVAMER 420
Db 377 GYFLYTSKSTFIINWTRVDSMDNDAYLGAAAIILNRVQSLVGLGAPRYOHIGLVAMER 436
Qy 421 QNTGWMESNANVKGTOIGAYFGASLCSDVDVNGSTDLVLI GAPHYETRGQGVSVCP 480
Db 437 QNTGWMESNANVKGTOIGAYFGASLCSDVDVNGSTDLVLI GAPHYETRGQGVSVCP 496
Qy 481 PRGQARWQCDVAVLGEQCPWGRFGAALTVDVNGDKLTDVAIGAPCEEDNRGAVYLF 540
Db 497 PRGQARWQCDVAVLGEQCPWGRFGAALTVDVNGDKLTDVAIGAPCEEDNRGAVYLF 556
Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDTVGAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDTVGAQGHVLLRSQ 616
Qy 601 PVLRVKAIEMFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLRRCQIOSVVT 660
Db 617 PVLRVKAIEMFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLRRCQIOSVVT 676
Qy 661 YDLALDSGRPHSRVAFNETKNSTRRTQVLTGTCTETLKLQLPNCIEDPVSPIVLRNF 720
Db 677 YDLALDSGRPHSRVAFNETKNSTRRTQVLTGTCTETLKLQLPNCIEDPVSPIVLRNF 736
Qy 721 SLVGTPLSAFNLRPVLAEDAORLTALPPFEKNCNDNICODDLSITFSFMSLDCLVWG 780
Db 737 SLVGTPLSAFNLRPVLAEDAORLTALPPFEKNCNDNICODDLSITFSFMSLDCLVWG 796
Qy 781 GPREFNVTVVRNDGEDSVYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 840
Db 797 GPREFNVTVVRNDGEDSVYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 856
Qy 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGKLLKANKYTSNNMPTNKTEF 900

Db 857 SGALKSTSCSINHPIFFENSEVTNITFOVDSKASLGNKLLKXANTSENMPRTNKTEF 916
Qy 901 QLELPVKYAVYVWVTSHGVSSTKYLNTASNTSRVMOHQYOVNSLGRSLPISLVLVPV 960
Db 917 QLELPVKYAVYVWVTSHGVSSTKYLNTASNTSRVMOHQYOVNSLGRSLPISLVLVPV 976
Qy 961 RLNQTVIWRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCDIP 1020
Db 977 RLNQTVIWRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCDIP 1036
Qy 1021 PFGIOEEFNATLKNLSFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGOGAFVRSQTEF 1080
Db 1037 PFGIOEEFNATLKNLSFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGOGAFVRSQTEF 1096
Qy 1081 KVEPFEVNPVPLPIVSGSVGLLLALITAAALYKLGFFKQYKDMMSGEGPPGABPQ 1137
Db 1097 KVEPFEVNPVPLPIVSGSVGLLLALITAAALYKLGFFKQYKDMMSGEGPPGABPQ 1153

RESULT 11

US-08-476-062A-43
Sequence 43, Application US/08476062A
Patent No. 5877275

GENERAL INFORMATION:

APPLICANT: Armaout, M. Amin

TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY

TITLE OF INVENTION: RESPONSES WITH BETA2 INTEGRINS

NUMBER OF SEQUENCES: 53

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.

STREET: 225 Franklin Street

CITY: Boston

STATE: MA

COUNTRY: US

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: Windows95

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/476,062A

FILING DATE: 07-JUN-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/216,081

FILING DATE: 21-MAR-1994

APPLICATION NUMBER: 07/637,830

FILING DATE: 04-JAN-1991

APPLICATION NUMBER: 07/539,842

FILING DATE: 18-JUN-1990

APPLICATION NUMBER: 07/212,573

FILING DATE: 28-JUN-1988

ATTORNEY/AGENT INFORMATION:

NAME: Freeman, John W.

REGISTRATION NUMBER: 29,066

REFERENCE/DOCKET NUMBER: 00786/068003

TELEPHONE: 617/542-5070

TELEFAX: 617/542-8906

TELEX: 200154

INFORMATION FOR SEQ ID NO: 43:

SEQUENCE CHARACTERISTICS:

LENGTH: 1152 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

FRAGMENT TYPE: internal

US-08-476-062A-43

Query Match 98.8%; Score 5808.5; DB 2; Length 1152;

Best Local Similarity 98.6%; Pred. No. 0;

Matches 1121; Conservative 8; Mismatches 7; Indels 1; Gaps 1;

Qy 1 PNLDTENAMTFOENARGFGQSVQLOGSRVVVGAPQETVAANORGSLYQCDYSTGSCBPI 60
Db 17 PNLDTENAMTFOENARGFGQSVQLOGSRVVVGAPQETVAANORGSLYQCDYSTGSCBPI 76
Qy 61 RLQVPEAVNNSLGLSLAATTSPOLLACGPTVHOTCSNTYVYKGLCFGLFNSLQQPQK 120
Db 77 RLQVPEAVNNSLGLSLAATTSPOLLACGPTVHOTCSNTYVYKGLCFGLFNSLQQPQK 136
Qy 121 PFEALRGCPQEDSDIAFLIDGSGSIIPDPRMKGELVSTIMEQLKSKTLPFLMOYSBEP 180
Db 137 PFEALRGCPQEDSDIAFLIDGSGSIIPDPRMKEFVSTVMEQLKSKTLPFLMOYSBEP 196
Qy 181 RHFTFKFQNNPNSRLIKPTQLGLGTHATGLRKVVRELFNITNGARKNAKILFLL 240
Db 197 RHFTFKFQNNPNSRLIKPTQLGLGTHATGLRKVVRELFNITNGARKNAKILFLL 256
Qy 241 TDGERFGDPLGYEDYIPELDRGVTIRYVGLFGDAPFRSEKSOBELMTVASKPRDHFVQAN 300
Db 257 TDGERFGDPLGYEDYIPEADREGVTIRYVIGVDAPFRSEKSOBELMTIASKPRDHFVQAN 316
Qy 301 NFEALKTVQNLREKI PALEGTCGSSSFEHMSOEGFSAAITNSGPIILSTVGSYDWAQ 360
Db 317 NFEALKTVQNLREKI PALEGTCGSSSFEHMSOEGFSAAITNSGPIILSTVGSYDWAQ 376
Qy 361 GYFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAIILNRVOSLVLGAPRYQHIGLVAMPR 420
Db 377 GYFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAIILNRVOSLVLGAPRYQHIGLVAMPR 436
Qy 421 QNTGMWESNANVKGTOIGAYFCASICSDVDVNSGSDTLVLCAPHYVQTRGGQSVVCLP 480
Db 437 QNTGMWESNANVKGTOIGAYFCASICSDVDVNSGSDTLVLCAPHYVQTRGGQSVVCLP 496
Qy 481 PRGQBARWOCDAVLYCEGQGPWGRFGAALTIVLGVNGDKLTDVAICAPGEEDNRGAVYLF 540
Db 497 PRG-QARWOCDAVLYCEGQGPWGRFGAALTIVLGVNGDKLTDVAICAPGEEDNRGAVYLF 555
Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGOSLGGODLTMDGLVLTWCAQGHVLLRSQ 600
Db 556 HGTSGSGISPSHSQRIAGSKLSPRLQYFGOSLGGODLTMDGLVLTWCAQGHVLLRSQ 615
Qy 601 PVLRVKAIKMFNPREVARNVFECNQVWKGKAGEVRVCLRVQKSTRDLREGQIQSVVT 660
Db 616 PVLRVKAIKMFNPREVARNVFECNQVWKGKAGEVRVCLRVQKSTRDLREGQIQSVVT 675
Qy 661 YDLALDSGRPHSRAVFNETKSTRQTVLGLTQTCETLKLQIPNCIEDVPSPVLRNLP 720
Db 676 YDLALDSGRPHSRAVFNETKSTRQTVLGLTQTCETLKLQIPNCIEDVPSPVLRNLP 735
Qy 721 SLVGTPLSAFGLNLRPVLAEADQRLFTALPPEKNCNDNI CODDLSITSPMSLDCLVVG 780
Db 736 SLVGTPLSAFGLNLRPVLAEADQRLFTALPPEKNCNDNI CODDLSITSPMSLDCLVVG 795
Qy 781 GPREFNVTVVRNDCGDSYRTQVTFPFLDLVYRVKSTLQNRQSRWSRLACESASSTEV 840
Db 796 GPRESNVTVVRNDCGDSYRTQVTFPFLDLVYRVKSTLQNRQSRWSRLACESASSTEV 855
Qy 841 SGALKSTSCSINHPIFFENSEVTNITPDVDSKASLGNKLLKXANTSENMPRTNKTEF 900
Db 856 SGALKSTSCSINHPIFFENSEVTNITPDVDSKASLGNKLLKXANTSENMPRTNKTEF 915
Qy 901 QLELPVKYAVYVWVTSHGVSSTKYLNTASNTSRVMOHQYOVNSLGRSLPISLVLVPV 960
Db 916 QLELPVKYAVYVWVTSHGVSSTKYLNTASNTSRVMOHQYOVNSLGRSLPISLVLVPV 975
Qy 961 RLNQTVIWRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCDIP 1020
Db 976 RLNQTVIWRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCDIP 1035
Qy 1021 PFGIOEEFNATLKNLSFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGOGAFVRSQTEF 1080
Db 1036 PFGIOEEFNATLKNLSFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGOGAFVRSQTEF 1095
Qy 1081 KVEPFEVNPVPLPIVSGSVGLLLALITAAALYKLGFFKQYKDMMSGEGPPGABPQ 1137

Db 1096 KVEPFEVNPPLPIVGVSSVGGILLALITAAALYKLGFFKQYKDMWSEGGPPGABPQ 1152
RESULT 12
PCT-US96-01314-43
; Sequence 43, Application PC/TUS9601314
; GENERAL INFORMATION:
; APPLICANT: M. Amin Arnaout
; TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN
; TITLE OF INVENTION: ANTAGONISTS
; NUMBER OF SEQUENCES: 78
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/01314
; FILING DATE: 30-JAN-96
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/380,167
; FILING DATE: 30-JAN-95
; ATTORNEY/AGENT INFORMATION:
; NAME: John W. Freeman
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00786/267001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1152
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
PCT-US96-01314-43
Query Match 98.8%; Score 5808.5; DB 5; Length 1152;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 1121; Conservative 8; Mismatches 7; Indels 1; Gaps 1;
Qy 1 FNLDTENAMTFQENARGFGQSVVQLOGSRVVGAPQEIIVAANQSGIYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFQENARGFGQSVVQLOGSRVVGAPQEIIVAANQSGIYQCDYSTGSCPEI 76
Qy 61 RLOQVPEAVNMSLGLSLAATTSPPQLLACQPTVHQTCSENYVYKGLCFPLFGSNLRQPOK 120
Db 77 RLOQVPEAVNMSLGLSLAATTSPPQLLACQPTVHQTCSENYVYKGLCFPLFGSNLRQPOK 136
Qy 121 FPALRGCPQEDSDIAFLIDGSGIIPHDPRMKELVSTIMEQLKSKTLPFLMQLYSEEP 180
Db 137 FPALRGCPQEDSDIAFLIDGSGIIPHDPRMKELVSTIMEQLKSKTLPFLMQLYSEEP 196
Qy 181 RIHPTFKFQNNPNRSLKIPITQLGRTHATGLRKRVRELNIINGARKNAKFLIL 240
Db 197 RIHPTFKFQNNPNRSLKIPITQLGRTHATGLRKRVRELNIINGARKNAKFLIL 256
Qy 241 TDGEKFGDPLGYEDVPELDREGVIRYVLGFGDAFSEKSRQELNTVASKPRDHDVFOAN 300
Db 257 TDGEKFGDPLGYEDVPELDREGVIRYVLGFGDAFSEKSRQELNTVASKPRDHDVFOAN 316
Qy 301 NFEALTKVQQLREKIPALIEGTQTGSSSPFHEMSQEGFSAALITNSGILLSTVGSYDWAG 360
Db 317 NFEALTKVQQLREKIPALIEGTQTGSSSPFHEMSQEGFSAALITNSGILLSTVGSYDWAG 376

Qy 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 436
Qy 421 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVIGAPHYBQTRGQSVCP 480
Db 437 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVIGAPHYBQTRGQSVCP 496
Qy 481 PRGQARWQCDVAVLGEQGPWGRFGAALTVDLGVNGDKLTDVAICAGEEDNRCGAVYLF 540
Db 497 PRG-PRWQCDVAVLGEQGPWGRFGAALTVDLGVNGDKLTDVAICAGEEDNRCGAVYLF 555
Qy 541 HGTSGSISPSHSQRIAGSKLSPLOYFGQSLSGQDLTMDGLVLTGVAQGHVLLRSQ 600
Db 556 HGTSGSISPSHSQRIAGSKLSPLOYFGQSLSGQDLTMDGLVLTGVAQGHVLLRSQ 615
Qy 601 FVLRVKAIWEZNPREVARNVFECDNDQVVKGEAGVRVCLHVQKSTRDLREGQIQSVVT 660
Db 616 FVLRVKAIWEZNPREVARNVFECDNDQVVKGEAGVRVCLHVQKSTRDLREGQIQSVVT 675
Qy 661 YDLALDSGRPHSRVAFNETKSTRTQTVLGLTQTCETLKLQLPNCIEDPVSPVILRLNF 720
Db 676 YDLALDSGRPHSRVAFNETKSTRTQTVLGLTQTCETLKLQLPNCIEDPVSPVILRLNF 735
Qy 721 SLVGTPLSAGNLRPVLAEDQRLFTALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 780
Db 736 SLVGTPLSAGNLRPVLAEDQRLFTALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 795
Qy 781 GPREENVTVTVRNDEGDSYRTQVTPFPPLDLSYRKVSTLONQSRQSRWLACESASTEV 840
Db 796 GPREENVTVTVRNDEGDSYRTQVTPFPPLDLSYRKVSTLONQSRQSRWLACESASTEV 855
Qy 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNPRNKTEF 900
Db 856 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNPRNKTEF 915
Qy 901 QLELPVKYAVVAVVTSHGVSITKYLNTASENTSRVMQHQYQVSNLQORSPLTSLVFLVEV 960
Db 916 QLELPVKYAVVAVVTSHGVSITKYLNTASENTSRVMQHQYQVSNLQORSPLTSLVFLVEV 975
Qy 961 RLNQTVIWDPRQVTFESNLSSTCHTKERLPSHSDPLAELKAPVNCSTAVCORIQCDIP 1020
Db 976 RLNQTVIWDPRQVTFESNLSSTCHTKERLPSHSDPLAELKAPVNCSTAVCORIQCDIP 1035
Qy 1021 PFGIOEEFNATLKNLSFDWYIKTSHNHLIIVSTABILFNDSVFTLLPQOGAFVRSQIET 1080
Db 1036 PFGIOEEFNATLKNLSFDWYIKTSHNHLIIVSTABILFNDSVFTLLPQOGAFVRSQIET 1095
Qy 1081 KVEPFEVNPPLPIVGVSSVGGILLALITAAALYKLGFFKQYKDMWSEGGPPGABPQ 1137
Db 1096 KVEPFEVNPPLPIVGVSSVGGILLALITAAALYKLGFFKQYKDMWSEGGPPGABPQ 1152
RESULT 13
5424399-2
; Patent No. 5424399
; APPLICANT: ARNAOUT, M. AMIN
; TITLE OF INVENTION: HUMAN CR3a/b HETERODIMERS
; NUMBER OF SEQUENCES: 12
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/78,871
; FILING DATE: 16-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 539,842
; FILING DATE: 18-JUN-1990
; APPLICATION NUMBER: 212,573
; FILING DATE: 28-JUN-1988
; SEQ ID NO: 2
; LENGTH: 1152
5424399-2
Query Match 98.8%; Score 5808.5; DB 6; Length 1152;

Best Local Similarity 98.6%; Pred. No. 0;
Matches 1121; Conservative 8; Mismatches 7; Indels 1; Gaps 1;

Qy 1 FNLDTENAMTFOENARGQSVVQLQGSRRVVGAPQOEIVAAQNGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFOENARGQSVVQLQGSRRVVGAPQOEIVAAQNGSLYQCDYSTGSCPEI 76

Qy 61 RLQVPVEAVNMISGLSLAATTSPPOLLACGPTVHQTCSNTYVKGCLFPLGSLNRQOPQK 120
Db 77 RLQVPVEAVNMISGLSLAATTSPPOLLACGPTVHQTCSNTYVKGCLFPLGSLNRQOPQK 136

Qy 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKELVSTIMEQLKSKTLFSLMOYSEEF 180
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKELVSTIMEQLKSKTLFSLMOYSEEF 196

Qy 181 RIHFTKEFQNNPRSLIKPITQLGRTHATGLRKVVRELFTNGARKNAFKILFLL 240
Db 197 RIHFTKEFQNNPRSLIKPITQLGRTHATGLRKVVRELFTNGARKNAFKILFLL 256

Qy 241 TDGKFGDPLGYEDVPELDRGVIRVYVVGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
Db 257 TDGKFGDPLGYEDVPELDRGVIRVYVVGDAFRSEKSRQELNTVASKPPRDHVFQAN 316

Qy 301 NFBAKTVQNLREKIFAIEGTQGSSTSPHEMSQSGFSAITNSGPELLSTVGSYDWAG 360
Db 317 NFBAKTVQNLREKIFAIEGTQGSSTSPHEMSQSGFSAITNSGPELLSTVGSYDWAG 376

Qy 361 GVFLYTSKXSTFTNMTRVDSMDMDAYLGYAAAIILNRVOSLVLGAPRYOHIGLVAMFR 420
Db 377 GVFLYTSKXSTFTNMTRVDSMDMDAYLGYAAAIILNRVOSLVLGAPRYOHIGLVAMFR 436

Qy 421 QNTGWMSNANVKTQIGAYGASLCSVDVDSNGSTDLVLIGAPHYETQTRGGQVSCPL 480
Db 437 QNTGWMSNANVKTQIGAYGASLCSVDVDSNGSTDLVLIGAPHYETQTRGGQVSCPL 496

Qy 481 PRGARWQCDAYLVGEGQDPWGFAGALTVLGVNKGDKLTDVAIGAPGEDNKGAYLFL 540
Db 497 PRGARWQCDAYLVGEGQDPWGFAGALTVLGVNKGDKLTDVAIGAPGEDNKGAYLFL 555

Qy 541 HGTSGSGISPHSQRISAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTVGAGHVILLRSQ 600
Db 556 HGTSGSGISPHSQRISAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTVGAGHVILLRSQ 615

Qy 601 PVLRVKATMEFNPREVANVFCNDQVVKGEAGBVRVCLAVQKSTRDRLEGQIQSVVT 660
Db 616 PVLRVKATMEFNPREVANVFCNDQVVKGEAGBVRVCLAVQKSTRDRLEGQIQSVVT 675

Qy 661 YDLALDSGRPHSRVAFNETKNSRRQTVGLGLTQTCETLKLQPNCTEDPVSPIVLRINF 720
Db 676 YDLALDSGRPHSRVAFNETKNSRRQTVGLGLTQTCETLKLQPNCTEDPVSPIVLRINF 735

Qy 721 SLVGTPLSAGNLRPVLAEDQRLFTALPPEKNGNDNICODDLSITFSPMSLDCLVVG 780
Db 736 SLVGTPLSAGNLRPVLAEDQRLFTALPPEKNGNDNICODDLSITFSPMSLDCLVVG 795

Qy 781 GPRFNVTVTVVRNDEGDSYRQVTFPPFLDLSYKSVSTLQNRQSRWRLACESASSTEV 840
Db 796 GPRFNVTVTVVRNDEGDSYRQVTFPPFLDLSYKSVSTLQNRQSRWRLACESASSTEV 855

Qy 841 SCALKSTSCSNHPFPENSVTNIIPDVDSKASLGNKLLKANVTSENMPNTKTEP 900
Db 856 SCALKSTSCSNHPFPENSVTNIIPDVDSKASLGNKLLKANVTSENMPNTKTEP 915

Qy 901 QLELPVKAVVMTVSHGVSTKYLNFNTASENTSRVMOHQYQVSNLQORSILPISLVFLVPV 960
Db 916 QLELPVKAVVMTVSHGVSTKYLNFNTASENTSRVMOHQYQVSNLQORSILPISLVFLVPV 975

Qy 961 RLNQTVIWRDQVTFSENLSTCTTKERLPKSHSDPLAELRKAPVWNCISVQRIQCDIP 1020
Db 976 RLNQTVIWRDQVTFSENLSTCTTKERLPKSHSDPLAELRKAPVWNCISVQRIQCDIP 1035

Qy 1021 FFGIOEEFNATLKGSLSPDWYIKTSHNLLIVSTABILFNDSVFTLLPQCGAFVRSQTET 1080

Db 1036 FFGIOEEFNATLKGSLSPDWYIKTSHNLLIVSTABILFNDSVFTLLPQCGAFVRSQTET 1095

Qy 1081 KVPEFPVPNPLPLIVGSSVGGLLLLALITAAALYKLGFPFRQYKDMSEGGPPGAEQ 1137
Db 1096 KVPEFPVPNPLPLIVGSSVGGLLLLALITAAALYKLGFPFRQYKDMSEGGPPGAEQ 1152

RESULT 14
US-08-476-062A-44
; Sequence 44, Application US/08476062A
; Patent No. 5877275
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY
; TITLE OF INVENTION: RESPONSES WITH BETA2 INTEGRINS
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,062A
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/216,081
; FILING DATE: 21-MAR-1994
; APPLICATION NUMBER: 07/637,830
; FILING DATE: 04-JAN-1991
; APPLICATION NUMBER: 07/539,842
; FILING DATE: 18-JUN-1990
; APPLICATION NUMBER: 07/212,573
; FILING DATE: 28-JUN-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Freeman, John W.
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00786/068003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1163 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-476-062A-44

Query Match 58.6%; Score 3446; DB 2; Length 1163;
Best Local Similarity 60.5%; Pred. No. 4.4e-278;
Matches 683; Conservative 143; Mismatches 297; Indels 6; Gaps 4;

Qy 1 FNLDTENAMTFOENARGQSVVQLQGSRRVVGAPQOEIVAAQNGSLYQCDYSTGSCPEI 60
Db 20 FNLDTENAMTFOENARGQSVVQLQGSRRVVGAPQOEIVAAQNGSLYQCDYSTGSCPEI 79

Qy 61 RLQVPVEAVNMISGLSLAATTSPPOLLACGPTVHQTCSNTYVKGCLFPLGSLNRQOPQK 120
Db 80 RLQVPVEAVNMISGLSLAATTSPPOLLACGPTVHQTCSNTYVKGCLFPLGSLNRQOPQK 137

Qy 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKELVSTIMEQLKSKTLFSLMOYSEEF 180
Db 136 LPVSRQECRQEQDIVFLIDGSGSIISRNFRATMNFRAVISQFRPSTQFSLMQFSNKF 197

Qy 181 RIHFTKEFQNNPRSLIKPITQLGRTHATGLRKVVRELFTNGARKNAFKILFLL 240

198 QTHFTFEERRTSNPLSLLASVHQLQGFYTTATAIQNVVHRLPHASYGARRDATKILIVI 257
241 TDGKFGDPLGYEDVTPELDREGVIRYVLGDFADRESEKSRQELNTVASKPDRDHVFOAN 300
258 TDGKFGDPLGYEDVTPELDREGVIRYVLGDFADRESEKSRQELNTVASKPDRDHVFOAN 317
301 NFEALKTQVONLREKIFAIEGTQTSSTSSFEHMSQGFSAATISNGPILSTVGSYDWAQ 360
318 DEFALKDQIQNOLKEKIFAIEGTQTSSTSSFEHMSQGFSAATISNGPILSTVGSYDWAQ 377
361 GVFLYTSKEKSTFNNTRVDSNDADAYLGYAAAILRNVRQSLVGLGAPRYOHTGAVIF 420
378 GAFLYPPNKSPTFNNQENVDMDSDYLGSTELALWKGVSQSLVGLGAPRYOHTGAVIF 437
421 QNTGWSNANVKTQIGAYFGASLCSVDVDSNGSTDLVIGAPHYEQTRGGQVSCPL 480
438 QVSRQWKAETVGTQIGSYFGASLCSVDVDSNGSTDLVIGAPHYEQTRGGQVSCPL 497
481 PRGORARWQCDVLYGBOQGFPGFAGALTULGVNGBKLTVAIGAPBEENRGAVYLF 540
498 PRGWR-RMWCDVLYGBOQGFPGFAGALTULGVNGBKLTVAIGAPBEENRGAVYLF 556
541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGOSLGGQDLMGDLVLTGCAQGHVLLRSQ 600
557 HGVLFPSISPSHSQRIAGSKLSPRLQYFGOSLGGQDLMGDLVLTGCAQGHVLLRSQ 616
601 PVLVRKAIEMFNPREVARNVFECDQVVKGAGEVVRVCLHVOKSTROBLRGOTQSVVT 660
617 PVLWGVSNQFTPAEIPRFAFECEQVSEOTLVQSNICLYIDKSKNLLGSRDLQSSVT 676
661 YDLALDSGPHSRVAFNETKSTRRTQVLTGTCETLKLQPNICRDPVPIVLRNF 720
677 LBLALDPGLSPRAFOETKNSLSRVRVLGKHCENENLLPSCVEDSVTPIRLNF 736
721 SLVGTPLSAPGNLPRVLAEDQELFTALPPFKNGCNDNICODDLSITFSEMSLDCLVVG 780
737 TLVGPCLLAFRLPMLAALQRYFTASLPFEKNGCADHICQDMLGISFSPGLKSLVG 796
781 GRPEENVVTVRNGEDSVRTQVTPPPDLISVRKUSTLQNGESQSRWELACESASTEV 840
797 SMLNLAENVWVNDGEDSVTITSHHPAGLSRYVABEQKQGLRSLHLTCDSPVGV-- 854
841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASIGKLLKANKVTSNNMPTNKTEP 900
855 SQGTWSTSCRIHNLIFGGAQITFLATPDVSPKAVLGDRLLTANVSSNNTPRTSKITF 914
901 QLELPVKYAVVWVTSYHGVSTYKLNFTAS-ENTSRVMOHQYVSNLGOBSLPISLVPLVP 959
915 QLELPVKYAVVWVTSYHGVSTYKLNFTAS-ENTSRVMOHQYVSNLGOBSLPISLVPLVP 974
960 VRLNOTVIWDRPQVTFSENLSTCHTKERLPSSDPLAELRKAPVNVCSIAVCORIQCDI 1019
975 VELNQAENVMDVEVSHFPQNPRLCSSEKIAPASDPLAHIQKNPVLDCSIAGCLPRCDV 1034
1020 PFFGIQEFNATLKNLSPDWYKITSNHLITVSTAEILFNDVSTFLIPGQAFVRSQTE 1079
1035 PSFSVQELDFTLKGNLSFGWRQILQKKVSVVVAEITFDTSVYSQLPQGEAFWEAQT 1094
1080 TKVBEPEVNPPLVGVSSVGLLALLALTYLKGPFKRYKDMWSE 1128
1095 TVLEKTKVHNTPPLVGVSSIGGLLALLATVLYKVGFKRYKDMWSE 1143

RESULT 15
PCT-US96-01314-44
; Sequence 44, Application PC/TUS9601314
; GENERAL INFORMATION:
; APPLICANT: M. Amin Attacut
; TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN
; TITLE OF INVENTION: ANTAGONISTS
; NUMBER OF SEQUENCES: 78
; CORRESPONDENCE ADDRESS:

ADDRESS: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 555X
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/01314
FILING DATE: 30-JAN-96
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/380,167
FILING DATE: 30-JAN-95
ATTORNEY/AGENT INFORMATION:
NAME: John W. Freeman
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 00786/267001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 1163
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
PCT-US96-01314-44

Query Match 58.6%; Score 3446; DB 5; Length 1163;
Best Local Similarity 60.5%; Pred. No. 4.4e-278;
Matches 683; Conservative 143; Mismatches 297; Indels 6; Gaps 4;

QY 1 FNLDTENAMTQENARGFGQSVWOLQGSRVVVGAPQEIIVAAQNGSLYQCDYSGSCPEI 60
DB 20 FNLDTBELTARVDSAGFGDSVQVYANSVWVGAPQKITANQGTGLYQCGYSTGACEPI 79
QY 61 RLQVPEAVNMSLGLSLAATSPPELLACGTVHQCSTENTYVGLCPLFCNLRQOQOK 120
DB 80 GLQVPEAVNMSLGLSLAATSPPELLACGTVHQCSTENTYVGLCPLFCNLRQOQOK 137
QY 121 FPEALRGCPQSDSIAPFLIDGSGSIIIPDRFMKELVSTIMEQLKSKTFLSLMOYSEEF 180
DB 138 LPVSRQECPRQDIIVFLIDGSGSISRNFAFMNFRAVISQPRSTQSLQFQSNKP 197
QY 181 RIHFTPEFQNNPNRSLIKPITQLGRHTHTATGLRVKVRFLNITNGARKNAFKILPLL 240
DB 198 QTHFTFEERRTSNPLSLLASVHQLQGFYTTATAIQNVVHRLPHASYGARRDATKILIVI 257
QY 241 TDGKFGDPLGYEDVTPELDREGVIRYVLGDFADRESEKSRQELNTVASKPDRDHVFOAN 300
DB 258 TDGKFGDPLGYEDVTPELDREGVIRYVLGDFADRESEKSRQELNTVASKPDRDHVFOAN 317
QY 301 NFEALKTQVONLREKIFAIEGTQTSSTSSFEHMSQGFSAATISNGPILSTVGSYDWAQ 360
DB 318 DEFALKDQIQNOLKEKIFAIEGTQTSSTSSFEHMSQGFSAATISNGPILSTVGSYDWAQ 377
QY 361 GVFLYTSKEKSTFNNTRVDSNDADAYLGYAAAILRNVRQSLVGLGAPRYOHTGAVIF 420
DB 378 GAFLYPPNKSPTFNNQENVDMDSDYLGSTELALWKGVSQSLVGLGAPRYOHTGAVIF 437
QY 421 QNTGWSNANVKTQIGAYFGASLCSVDVDSNGSTDLVIGAPHYEQTRGGQVSCPL 480
DB 438 QVSRQWKAETVGTQIGSYFGASLCSVDVDSNGSTDLVIGAPHYEQTRGGQVSCPL 497
QY 481 PRGORARWQCDVLYGBOQGFPGFAGALTULGVNGBKLTVAIGAPBEENRGAVYLF 540
DB 498 PRGWR-RMWCDVLYGBOQGFPGFAGALTULGVNGBKLTVAIGAPBEENRGAVYLF 556

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 17:05:16 ; Search time 14.559 Seconds
(without alignments)
7512.163 Million cell updates/sec

Title: US-09-902-481B-4
Perfect score: 5884
Sequence: 1 FNLDENAMTQENARGFGQ.....FKRQYKDMWSEGGPRGAPQ 1137

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*

1: p1r1:*

2: p1r2:*

3: p1r3:*

4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5852	99.5	1153	1 RWHUIB	cell surface glyco
2	4460	75.8	1153	2 S00551	leukocyte surface
3	3469	59.0	1163	1 RWHUIC	cell surface glyco
4	1543.5	26.2	1170	2 S03308	cell surface glyco
5	1527.5	26.0	1163	2 I56126	lymphocyte fuction
6	1140	19.4	1179	2 A53213	integrin alpha-E c
7	1093.5	18.6	1151	2 A45226	integrin alpha-1 c
8	1071	18.2	1170	2 I45914	integrin alpha 2 s
9	1057	18.0	1178	2 S44142	VLA-2 protein homo
10	1055	17.9	1180	2 A35854	integrin alpha-1 c
11	1054	17.9	1181	2 A33998	integrin alpha-2 c
12	663	11.3	1039	2 R41131	lymphocyte-Peyer's
13	642	10.9	1038	2 S06046	integrin alpha-4 c
14	630	10.7	1035	2 I58409	integrin alpha-9 c
15	614.5	10.4	1041	2 T31437	integrin alpha cha
16	579.5	9.8	1054	2 JC7294	alphanap integrin -
17	572.5	9.7	1051	2 A35761	cell surface glyco
18	567.5	9.6	1053	2 I55534	VLA-3 alpha subuni
19	555.5	9.4	1053	2 S44250	integrin alpha-5 c
20	546.5	9.3	1034	2 A36108	integrin alpha-V c
21	535	9.1	1044	2 T10050	integrin alpha-v c
22	534.5	9.1	1072	2 A38457	integrin alpha-6 c
23	532	9.0	1049	2 A27079	fibronectin recept
24	532	9.0	1073	2 B36429	integrin alpha-6 c
25	530.5	9.0	1051	2 A40021	integrin alpha-3 c
26	526	8.9	1048	2 A27421	integrin alpha-5 c
27	523.5	8.9	1091	2 A41543	integrin alpha-6 c
28	517	8.8	1044	2 S16516	integrin alpha-8 c
29	496.5	8.4	1146	2 S40311	integrin - fruit f

ALIGNMENTS

RESULT 1

RWHUIB

cell surface glycoprotein CD11b precursor [validated] - human

N:Alternate names: complement receptor type 3 alpha chain; leukocyte adhesion protein M
eukocyte integrin alpha chain; neutrophil adherence receptor alphaM chain

C:Species: Homo sapiens (man)

C>Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text change 31-Dec-2000

C:Accession: A31108; A28915; A41600; A30892; A32218; A46526; A26091; I52567

R:Corbi, A.L.; Kishimoto, T.K.; Miller, L.J.; Springer, T.A.

J. Biol. Chem. 263, 12403-12411, 1988

A:Title: The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, C
B.

A:Reference number: A31108; MUID:88315033; PMID:2457584

A:Accession: A31108

A:Molecule type: mRNA

A:Residues: 1-1153 <COR>

A:Cross-references: GB:J03925; NID:g187284; PIDN:AA59544.1; PID:g307148

A>Note: part of this sequence was confirmed by protein sequencing

R:Arnaout, M.A.; Gupta, S.K.; Pierce, M.W.; Tenen, D.G.

J. Cell Biol. 106, 2153-2158, 1988

A:Title: Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor 1

A:Reference number: A28915; MUID:88257215; PMID:2454931

A:Accession: A28915

A:Molecule type: mRNA

A:Residues: 1-499,501-965, 'P', 967-1153 <ARN>

A:Cross-references: GB:M18044; GB:J03270; GB:M19664; GB:X07421; NID:g186935; PIDN:AA59

A>Note: The authors translated the codon TAC for residue 1129 as Thr

A>Note: part of this sequence, including the amino end of the mature protein, was confi

R:Shelley, C.S.; Arnaout, M.A.

Proc. Natl. Acad. Sci. U.S.A. 88, 10525-10529, 1991

A:Title: The promoter of the CD11b gene directs myeloid-specific and developmentally re

A:Reference number: A41600; MUID:92073318; PMID:1683702

A:Accession: A41600

A:Molecule type: DNA

A:Residues: 1-9 <SHE>

A:Cross-references: GB:M76724; NID:g180018; PIDN:AA58410.1; PID:g553215

R:Arnaout, M.A.; Remold-O'Donnell, E.; Pierce, M.W.; Harris, P.; Tenen, D.G.

Proc. Natl. Acad. Sci. U.S.A. 85, 2776-2780, 1988

A:Title: Molecular cloning of the alpha-subunit of human and guinea pig leukocyte adhes

A:Reference number: A94193; MUID:88190151; PMID:2833753

A:Accession: A30892

A:Molecule type: mRNA

A:Residues: 917-1042 <AR2>

A:Cross-references: GB:M18044

R:Hickstein, D.D.; Hickey, M.J.; Ozols, J.; Baker, D.M.; Back, A.L.; Roth, G.J.

Proc. Natl. Acad. Sci. U.S.A. 86, 257-261, 1989

A:Title: cDNA sequence for the alphaM subunit of the human neutrophil adherence recepto

A:Reference number: A32218; MUID:89098893; PMID:2563162

A:Accession: A32218

A:Molecule type: mRNA

A:Residues: 9-1153 <HIC>

A:Cross-references: GB:J04145; NID:g189068; PIDN:AA59903.1; PID:g386975

R:Sastre, L.; Roman, J.M.; Teplow, D.B.; Dreyer, W.J.; Gee, C.E.; Larson, R.S.; Roberts, Proc. Natl. Acad. Sci. U.S.A. 83, 5644-5648, 1986
 A:Title: A partial genomic DNA clone for the alpha subunit of the mouse complement receptor
 A:Reference number: 159078; MUID:86287312; PMID:2942940
 A:Accession: 159078
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 11-44 <RES>
 A:Cross-references: GB:M14293; NID:g198993; PIDN:AAA39484.1; PID:g554193
 C:Genetics:
 A:Gene: Mac-1
 C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homology
 C:Keywords: cell adhesion; glycoprotein; transmembrane protein
 F:1-16/Domain: signal sequence #status predicted <SIG>
 F:17-1153/Product: leukocyte surface glycoprotein Mac-1 alpha chain #status experimental
 F:148-319/Domain: von Willebrand factor type A repeat homology <VWA2>
 F:1106-1129/Domain: transmembrane #status predicted <TM>

Query Match 75.8%; Score 4460; DB 2; Length 1153;
 Best Local Similarity 73.8%; Pred. No. 6e-304;
 Matches 840; Conservative 146; Mismatches 150; Indels 2; Gaps 2;

Qy	1	FNLDTENAMTPOENARGCOSVVOLOGSRVUVGAPQEIIVANQSGSLYOCDYSTGSCBPI	60
Db	17	FNLDTEHPMTPOENARGPQNNVQLGTSVVVAARQEAQVNGTGLAYQCDYSTSRCHPI	76
Qy	61	RLQVPEAVNMSLGLSLAATTSPPQLLACGPTVHQTGENTYVXGLCFGLFGLNLRQOPQK	120
Db	77	PLQVPEAVNMSLGLSLAVSTVPQQLACGPTVHQNCKENTVYVGLCYLFGSLNLRPQQ	136
Qy	121	FPALRGCPQESDIAFLDGSIIIPHPFRMKEVSVTMSQLKSKTLPFLMAYSEEF	180
Db	137	FPALRECPQESDIIVFLDGSININIDFQKMKFPVSTVMSQFKSKTLPFLMAYSEEF	196
Qy	181	RIHFTKPKFONNPRLSKPTQLGLRTHATGLKVVRELPENITNGARKNAKILFLL	240
Db	197	RIHFTFNDKPNPSHVSPIQLNGRTKTASGIRKVVRELPFKINGARENAKILVVI	256
Qy	241	TGKEFGDPLGYEDVTEIPDLREGVIRYVIGVGDAFRSEKSRQELNATVASKPRDHVQIN	300
Db	257	TGKEFGDPLDYKDVPEADRAGVIRYVIGVGNFKNPQSRRELDITASKPAGEHVQVD	316
Qy	301	NFALATIQNLQEKIPEAGTGTGSSGFHEMSQEGFSAAITSGPLISTVGSVDWAG	360
Db	317	NFALATIQNLQEKIPEAGTGTGSSGFHEMSQEGFSASITSGPLIGSVGSFDWAG	376
Qy	361	GVFLYTSKSKSTFINMTRVDSMDAYLGVAAIILRNVRQSLVTLGAPRYOHIGLVMPR	420
Db	377	GAFLYTSKDKVTINTRVDSMDAYLGVASAVILRNVRQSLVTLGAPRYOHIGLVMPR	436
Qy	421	QNTGMESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYVEOTRGQSVSCPL	480
Db	437	ENFGTWPHTSISKSGISYFGASLCSVDMDAGNTNLILIGAPHYKTRGQSVSCPL	496
Qy	481	PRQORARQCDVLYEGQGFWRFGAALTVDGVNGDKLTDVAIGAPCEEDNRGAYILF	540
Db	497	PRG-RARQCEALLHGQGHFWFGAALTVDGVNGDKLTDVAIGAPCEQNGQAVYIF	555
Qy	541	HGTSGGISGSHSQRIAGSKSLPRLQYFGQSLGGQDLTMDGLVLTGCAQCHVLLRSQ	600
Db	556	YGASIASLSASHSHRIAGHFSFGLOQYFGQSLGGQDLTMDGLMDLAVGAQGHLLLRQ	615
Qy	601	PVLVRVAIMEFNPREVARNVFCNDQVKGKAGEVRVCLVHVKSTRDLRREGQIOSVVT	660
Db	616	PVLRLATEMFEPSKVAARSVFAQEQVLKXKDAAGEVRVCLVRKNTKRLREGDIQSTVI	675
Qy	661	YDLALDSGRPHSRVAVNETKNSRQTVGLQTCETLKLQPLNCIEDPVSIVLRINEF	720
Db	676	YDLALDPVRSIRAFDFTKNTTRRTQVFLQKCEKTELKLIPLDCVDSVSPHILRLNY	735
Qy	721	SLVCTPLSARGNRPVLAEDAQRLEFALPPEKNCNDNICODLSTTFSFMSLDCLVVG	780
Db	736	TLVGEPLRSFGNLRPVLAEDAQRLEFALPPEKNCNDNICODLSTTFSFMSLDCLVVG	795

781 GPREFNVTVTRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNSORSRWL-ACESASSTE 839
 796 GPQDFNVSTLRNDGEDSYGTQTVVYPSGLSYKDSASQNPITKPFVFKPAESSSSSE 855
 840 VSGALKSTSCSINHPIPPENSEVTENITPDVDSKASIGNKLLKANKVTSNNMPTNKTE 899
 856 GHGALKSTTWNHPIPPANSEVTENITPDVDSKASIGNKLLKAIIVASNNMSTRHTKX 915
 900 FOELPLPKYAVYVTVTSHGVSTKYLAFTASENTSRVNHQYQVSNLQORSLPISLVFLVP 959
 916 FOELPLPKYALYVTVTSHGVSTKYLAFTASENTSRVNHQYQVSNLQORSLPISLVFWFIP 975
 960 VRLNQTVMWRPQVTFBENLSTCHTKERLPSHSDFLAELKAPVNCSTAVCORICODI 1019
 976 VQINNVTVMWPQVIFQNSLSSACHTEQSPHNSFRDQLERTPVLNCSVAVCKRIQCDL 1035
 1020 PFGIQEENFATLKNLSFDWYIKTSHNHLIYSTABILFNDSTVFTLLPGQGAFFVRSQTE 1079
 1036 PSFTQBIFFVTLKGNLSPDWYIKTSHGILLVSVSTELFNDSTAFALLPQGESVRSKTE 1095
 1080 TKVPPFVPPNPLPIVGVSSVGLLLALITAAALYKGLFFKQYKDMSEGGPPGAPQ 1137
 1096 TKVPPFVPPNPLPIVGVSSVGLLLALITAGLYKGLFFKQYKDMSEGGPPGAPQ 1153

RESULT 3
 RHEUIC
 cell surface glycoprotein CD11c precursor - human
 C:Alternate names: leukocyte adhesion receptor p150,95 alpha chain
 C:Species: Homo sapiens (man)
 C:Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text_change 22-Jun-1999
 C:Accession: A36584; A35543; S00864
 R:Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.
 J. Biol. Chem. 265, 12750-12751, 1990
 A:Reference number: A36584
 A:Contents: erratum
 A:Accession: A36584
 A:Molecule type: DNA
 A:Residues: 1-1163 <COR>
 A:Note: this revision to the sequence from reference A35543 includes the carboxyl end
 R:Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.
 J. Biol. Chem. 265, 2782-2788, 1990
 A:Title: Genomic structure of an integrin alpha subunit, the leukocyte p150,95 molecule.
 A:Reference number: A35543; MUID:90153906; PMID:2303426
 A:Accession: A35543
 A:Molecule type: DNA
 A:Residues: 1-834 <CO2>
 A:Note: this sequence has been revised in reference A36584
 R:Corbi, A.L.; Miller, L.J.; O'Connor, K.; Larson, R.S.; Springer, T.A.
 EMBO J. 6, 4023-4028, 1987
 A:Title: cDNA cloning and complete primary structure of the alpha subunit of a leukocyte
 A:Reference number: S00864; MUID:88166645; PMID:3327687
 A:Accession: S00864
 A:Molecule type: mRNA
 A:Residues: 1-755, 'L', 757-1163 <CO3>
 A:Cross-references: GB:M81695; EMBL:Y00093; NID:g487829; PIDN:AAA59180.1; PID:g487830
 A:Note: part of this sequence was confirmed by protein sequencing
 C:Comment: A common beta chain (CD18) forms a heterodimer with CD11b to form Mac-1 on my
 C:Genetics:
 A:Gene: ITGAX; CD11C
 A:Cross-references: GDB:119758; OMIM:151510
 A:Map position: 16p11.2-16p11.2
 C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homok
 C:Keywords: calcium; cell adhesion; glycoprotein; heterodimer; magnesium; tandem repeat,
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-1163/Product: cell surface glycoprotein CD11c #status predicted <MAT>
 F:20-1107/Domain: extracellular #status predicted <EXT>
 F:149-319/Domain: von Willebrand factor type A repeat homology <VWA4>
 F:1108-1163/Domain: transmembrane #status predicted <TM>
 F:1134-1163/Domain: intracellular #status predicted <INT>
 F:61,89,392,697,735,899,939,1050/Binding site: carbohydrate (Asn) (covalent) #status pre

Query Match	59.0%	Score 3459	DB 1	Length 1163
Best Local Similarity	60.9%	Pred. No. 1.8e-234		
Matches 688	Conservative 142	Mismatches 293	Indels 6	Gaps 4
QY	1	FNLDTENAMTFQGNARGFGQSVVQLOGSRVVVVGAPQIEIVAANQRGSLYQCDYSTGSCBPI	60	
DB	20	FNLDTBELTAFRVDSAGFGDSVVQYANSMVVVGAPQIKITAAQTGGLYQCGYSTGACBPI	79	
QY	61	RLQVPEAVNMSLGLSLAATTSPPOLLAQPTVHQTCSENTYVKGCLCFPLPGSNLRQPOK	120	
DB	80	GLQVPPPEAVNMSLGLSLASTTSPLLACQPTVHCEGRNMVLTGLCFLLGPT--QLTQR	137	
QY	121	FPBALRGCPQSDIAFLIDGSGIIPHDPFRMKEWVSTVMEOLKSKTLLFSLMXYSEHF	180	
DB	138	LPVSRGCPQEQODIVPLIDGSGISIRNPATWMPFRAVISQFQSPSTQSLAQESNKF	197	
QY	181	RIHFTPKFQNNPNSRLIKPIITOLLGRTHATGLAKRVVRELFININGARKNAFKILFL	240	
DB	198	QTHFTPEEFRTSNPLSLLASVHQLOQFTYTATAIQWVHRLPHASYGARRDATKILVI	257	
QY	241	TDCEKGDPLGYEDVTELPDREGVIRVVGVDAPFRSEKSRQSLNTVASKPPDRHVFQIN	300	
DB	258	TDGKKGDSLOYKDVIPMDADAGIIRTAIOGVGLAFQNRNSWKSENDIASKPSQEHIFKYE	317	
QY	301	NFEALKTIONQLREKIFAIBGTOTGSSSEFHEMSQEGFSAALITNSGPLLSTVSYDMA	360	
DB	318	DFDALKDIONQLREKIFAIBGTETSSSSFELEWAQEGPSAVTPDGPVLGAVGSPFWS	377	
QY	361	GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAILLRNVOSIVLGAPOYHGLGVAMFR	420	
DB	378	GAFLYPPNMSPTFTINMSQENVDNRDSYLGYSTELALWKGVOSIVLGAPOYHGTGKAVI	437	
QY	421	QNTGWNSSNANVKTQIGAVFGASLCSDVDVDSNGSTDLVLIGAPHYYEQTRGQGVSVCL	480	
DB	438	QVSRQWEMKAEVGTQIGSVYFGASLCSDVDVDTGSDTDLVLIGAPHYYEQTRGQGVSVCL	497	
QY	481	PRGORARWQCDVLYGBOQCPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVLYF	540	
DB	498	PRGWR-RWMDCAVLYGBOQCPWGRFGAALTVLGDVNGDKLTDVVI GAGPEENRGAVLYF	556	
QY	541	HGTSGSGISPSHSORIASGLSPRLQYFGQSLGGQDLTMDGLVDLTVGAQGHVLLARSQ	600	
DB	557	HGVLGPSISPSHSORIASGLSSRLQYFGQALSGQDLTQDGLVDLVAVGARGQVLLLR	616	
QY	601	PVLRVKALMEFNPREVARNYFECNDQVVGKKEAGEVRVCLHVKOKSTRDLREQIQSVVT	660	
DB	617	PVLWVGYSMQFIPAEIIPRSAPFECREQVVSQTLVQSNICILYDKRSNLLGSRDLQSSVT	676	
QY	661	YDLALDSGRPHSRAVENETGNSIFRQTVQVGLTQTCETLKLQPCNCTEDPVSPIVLRNLF	720	
DB	677	LDLALDPGLSPRATFOETKNRSUSRVVVLGLKAHCENFNLLPSCVEDSVITPILRLNF	736	
QY	721	SLVCTPLSAFONLAPVLAEDAQRLLFTALPPFKNCGNDNI CODGLSITTFMSLDCLVWG	780	
DB	737	TVNGKPLLAFLNAPMLAADAQRYFTASLPPFKNCGADHICQDNHGLSIFSPFPGLSLVG	796	
QY	781	GPRFENVTVVRNDGDSYRTQVTFPPPLDLSYRKVSTLQWRSQRWRWLRACASASTEV	840	
DB	797	SNLELNAEVMVWNDGDSYGTITFISHPAGLSYRYVAEGOKQQLSLHCTCDSPAGV--	854	
QY	841	SGALKSTSCSINHPIPFENSEVTENITFDVDSKASLGNKLLKANVTSENMMPTNKTEF	900	
DB	855	SGQWTASTSCRINHLLIPGGAQITFLATFDVSPKAVLGRLLLTANVSSENTPRTSKTTP	914	
QY	901	QLELPVKYAVTVMTVTSHGVSITYKLYNFTAS-ENTSRVMQHQYQVSNLQORSPLTSLVPLVP	959	
DB	915	QLELPVKYAVTVTVVSSHEQFTKYTLNFSESEKESHVAMHYQVNVNLRQORDLPVSNFMPV	974	
QY	960	VLRLNQTVWDRPOVTFPSNLSSTCHTKERLPSHSDPLAELRKAPVNVNCSIAVCORIQCDI	1019	
DB	975	VELNQEAVMVDVESHQNPDLRCSSEKIAPPASDFLIATQKNPVLDCSIAGCLRFECDV	1034	
QY	1020	PFFGIQEAFNATLKNLSFDWYIKTSHNHLILVSTASILFNDSVFTLLPQGFVRSQTE	1079	


```
QY 768 TFSFMSLDCLVGGP-----BFFNVTVTVRNDGDSYRTQVTFPPFLDLSYRKVSTLQW 821
Db 780 SSPARS-----GFLRMSASLAWETLNSGDAYVRLDLDPRGLSFKVEMLQ- 831
QY 822 QRSQRWRLACASSTEVSGAL-KSTSCSINHPIIPENSEVTNITFDVDSKASLGNKL 880
Db 832 --PHSRMPVSCBEL--TEGSLTLTKLKNVSPFKAGQGVSLQVNFNTLLNSSWEDFV 887
QY 881 LLKANTSEN-NMPTNKTFFOLELPVKYAVYVWVTSHTVSTKYLNFPTASENTSRVQHQ 939
Db 888 ELNGTVHCENENSLOQDSAAATHIPVLYPNVILTKQENSTLYISFTPKGPTQVQHV 947
QY 940 YQVSNLQGRSLPISLVFLVPRVRLNQTVIDRPO-----VTFSENLS--TCHTKE-RLP 990
Db 948 YQV-----RIQPAYDHNMT-LEALVGPVRPHSEDILITVMSVQTDPLVTCHEDLKRP 1001
QY 991 SHSDFLAELRKAPVNVCSIAVCORIQDIPFGIOBEFNATLKNLSFDWYIKTSHNLL 1050
Db 1002 SSE--AEQCLPGV-----QFRCPIVP---RWELLQVGTVELSKELKAS-STLS 1046
QY 1051 IVSTAELFNDVSFTLLPGQAFVRSQTETKVPPFVNPPLPIVIGSSVGLLLALITA 1110
Db 1047 LCSSLVSFNSSKHFHLYGSKA-SEAQVLVKVDLIEKEMLVVYVLSGIGGLVLLFLIFL 1105
QY 1111 ALYKLGEPKQYQDM-M-SEGPPGAP 1136
Db 1106 ALYKVGFFKRLKEKMEADGGVNGSP 1132

RESULT 6
A:Species: Homo sapiens (man)
C:Date: 19-Oct-1995 #sequence_revision 31-Mar-2001 #text_change 31-Mar-2001
C:Accession: A53213
R:Shaw, S.K.; Cepsek, K.L.; Murphy, E.A.; Russell, G.J.; Brenner, M.B.; Parker, C.M.
J. Biol. Chem. 269, 6016-6025, 1994
A:Title: Molecular cloning of the human mucosal lymphocyte integrin alpha (B) subunit.
A:Reference number: A53213; MUID:94164962; PMID:8119947
A:Accession: A53213
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1179 <SHA>
A:Cross-references: GB:L25851; NID:G457244; PID:G457245
A:Gene: GDB:ITGAE
A:Cross-references: GDB:330801
A:Map position: 17p13
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homology
F:199-371/Domain: von Willebrand factor type A repeat homology <WMA3>

Query Match 19.4%; Score 1140; DB 2; Length 1179;
Best Local Similarity 28.9%; Pred. No. 3.5e-71;
Matches 340; Conservative 214; Mismatches 451; Indels 172; Gaps 39;

QY 45 GSLYQCDYSTGS--CEPT-RLQVP-----VEAVMSLGLSLAATTPPOLLACGPVHQ 95
Db 65 GFLRCSLVQDEILCHPVEHPVPIPKGRHGVTVVRSHHGLICI-----QVLVRP--HS 117
QY 96 TCSENTYVKGCLPLFGSNLRQOPQ----- 119
Db 118 LSSELT---GTCSLGSLPDLRFOQANFPDENLDPDARVDVTGDCYSKSGGGEDDVNTA 174
QY 120 KFPBALRGCPQED-----SDIAFLIDGSGSIIPHDFRPMKEWVSTVWEOL-- 164
Db 175 RQRRALEKEEBEEDKEEBEEDBEAETAIALLDGSISIDPPDFQRAKDFISNMNRFFE 234
QY 165 KSKTSLSLMOYSEFRHFFPKFQNNPNRSLIKPTOLLGTHTHATGLRKVVRLFN 224
Db 235 KCFCNFAVQYGGVQTFEFLDRSDQVMSLARVQMTVGSVTKTASAMQHVLDISFT 294
QY 225 ITNGARKNAKILFLLTDEKFGDPLGYEDVIPEDREGVIRYVIGVDAPRSEKSKQEL 284
Db 225 ITNGARKNAKILFLLTDEKFGDPLGYEDVIPEDREGVIRYVIGVDAPRSEKSKQEL 284
```

```
Db 295 SSHGSRKASKVMVVLTDGIFEDFLNLTVTVINSPMQGVVERFAIGVGEFPKSAKTAREL 354
QY 285 NTVASKPRDRHVQINNFEALKTIOQLREKIFAIEGTQTGSSSSSEHEMSQEGSAAIT 344
Db 355 NLIASDPDETHAFKVTNWDALDGLSKLYNIIISMEGT---VGDALHYOLAQIGFSAQIL 411
QY 345 SNGP--LLSTVGSYDWAGGVFLY-TSKKSTTINMTRVDSMDND-----YLGAAAILLEN 398
Db 412 DERQVLLGAVGAFDWSGGALLYDTRSRGRFLNQTFAAAAADABAAQSYLYGVAVVLHKT 471
QY 399 RVQSILVIGAPRYOHLGLVAMER-QNTGMWESNANY-KGTQIGAYFGASLCSVDVDSNGST 456
Db 472 CSLSVTAGAPQKHEG--AVFELQKEGREASFPLVLEGEOMGSIYGSBELCPVIDMDGST 529
QY 457 DLVLICAPHYEQTGGQVSVCPPLPRQORARWQCDAVLYGSGQGPWGFAGALTVLGDVN 516
Db 530 DFLVAAAPFYHVHGEGRVYVYVRLSE-QDGSFSLARILISGHGFTNARFGFAMAAGDLS 588
QY 517 GPKLTDVAIGAP---GEEDNR--CAVYLFHGTSGGI:SPSHSQRISAGSKLSPLRQYFEQ 570
Db 589 QDKLTDVAIGAPLEGFGADDDGASFGSVIYNG-HWDGLSASPSQIRASTVAPGLQYFGM 647
QY 571 SLSGGQDLTMDGLVDLTVGAQGHVLLRSQPVLRVKALIMBNPRVARNVFECNDQVVG 630
Db 648 SWAGGFDISGDLADITVGTLCQAVVFRSRPVVRLKVSMAFTPSALP-----IGF 697
QY 631 KEAGSVVCLHVQKSTRDLREGQIQSVYVYDLDLSDGRPHSRVAFNETKNTSTRQTQVL 690
Db 698 NGVAVNRLCFEI-SSVTTASBSGLREALNFLVDVKGQRRLQCSVRSCLGCLREWS 756
QY 691 GLTQTCETLKLQLPN---CIEDPVSVILNLFSLVGTPLSAFGNLRPLVLAEDAQRLFT 746
Db 757 SSCQLCEDL-LAMPTEGELCEDCFSNASVKVSYQL-OTPEQQTQHPQILDRYTEPPAI 814
QY 747 ALPPEKNGNDNI CODDLISITFSFMSLDCLVGGPVRNVTVTVRNDGDSYRTQVTFP 806
Db 815 FOLPVEKACKNLCFVAELQLA-TTVSQELVVGUTKELTMINETNSGDSYMTSMALN 873
QY 807 FPLDLSYRKVSTLQORSORSLACESASSTEVSGALKSTSCSINHPIIPENSEVTNFI 866
Db 874 YPRNLQ-----LKMQRKPPSPNI QCDPPQPV---ASVLMNCRIGHPLV-KRSSAHVS 923
QY 867 TFDVDSKASLGNKLKILKANVTSENN-----MPTNKTBEQ---LELPKVAVYVWVTSHV 919
Db 924 VMQLSENAFPNRTADITVTVTNSNERRSLANETHLQFPHGFAVLKSPISMYVNTGQGL 983
QY 920 S--TKYLNFTASENTSRVMOHQYQVSNLQORSPLISLVFLVPRVRLNQTVIDRPOVTFSE 977
Db 984 SHHKEFLFHVHGEN---LFGASYQ-----LQICVPTKLRGLQVAAVXKLTRTQ 1028
QY 978 NLSSTCHTKERLPSSHSDFLAELRKAPVNVCSIAVCORIQDIPFGIOBEFNATLKNLS 1037
Db 1029 ASTVCTWSQERACAYSS-VQHVVEWHSVSCVIA-----SDKENVTVAABIS 1073
QY 1038 EDWYIKTSHNLLIVST-----AEILFNDVSFTLLPGQAFVRSQETKVPFEPVNPPL 1091
Db 1074 WD-----HSELLKDVTELOLIGELISFNKSLYEGNAENH--RTKITVTVFLKDKYHSL 1125
QY 1092 PLTVGSSVGLLLALITAAALYKLGFFKQYQDMME 1128
Db 1126 PIIIKSGVGLLVILVILFKCGFFKRYQQLALE 1162
```

RESULT 7

A45226
Integrin alpha-1 chain - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 30-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 15-Sep-2003
C:Accession: A45226
R:Briesewitz, R.; Bostein, M.R.; Marcantonio, E.E.
J. Biol. Chem. 268, 2999-2996, 1993
A:Title: Expression of native and truncated forms of the human integrin alpha 1 subunit
A:Reference number: A45226; MUID:93155124; PMID:8428973

A:Accession: A45226
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-1151 <BRI>
A:Experimental source: hepatoblastoma cell line HepG2
A>Note: sequence extracted from NCBI backbone (NCBIP.124326)
P:142-317/Domain: von Willebrand factor type A repeat homology <VWA1>

Query Match 18.6%; Score 1093.5; DB 2; Length 1151;
Best Local Similarity 27.8%; Pred. No. 6.1e-68;
Matches 345; Conservative 206; Mismatches 487; Indels 203; Gaps 44;

QY 1 FNLDTENAMTFQENARG-FQGSVVQL---QGSRRVVVGAPOEIVAAQNRGLSQCDYSGS 56
DB 1 FNVVKNKNTSGPVEDMGYTVQYENEEGKVLIGSLVGPQKRTGDVYKCPVGRGE 60
QY 57 CSP-IRLVQVRA-----VNSLGLSLAATTSPPOLLACGPTVHGTCSNTYVKG 106
DB 61 SLPCVXLDLFVNTSIENVTVEKNMTFGSTL-VTPNPGGLACGPLYAYRGHLYTGI 119
QY 107 CFLFGSLNQAPQKPEALRGCPQEDSDIAFLIDGSGSIIPDFRMEKWSVTWVQ 166
DB 120 CSDVSTFPQVNSIAP-VQECSTQ-LDIVVLDSNSIYPWDS-----VTAFLNDLLK 170
QY 167 -----SKTLFSLMYSSEFRHFTFKFQNNPNPSRLIKPTQLGR-THATGLRKVV 219
DB 171 RMDIGPKQTVQVIGYGENVTHEFNLNKYSSTEEVLVAAKIVQSGROTMALGTDYAR 230
QY 220 RELFNITNGARKNAFKILPILDTGEGFGLGYEDVPELDREGVIRYVIGYDAFR--- 276
DB 231 KEAFTARGARGKVKVWIVTDGESH-DNHRUKVIOCEDNIQRFIALIGSTNRYGN 289
QY 277 --SEKROELINTVASKPRPDHVPQINNFEALKTIONQREKIPAIPTGTQSGSSPFHEM 334
DB 290 LSTKEFVEEIKSTASEPTEKHFFNVSDDELAVIIVTGLGERIPALEATADQSAASP 349
QY 335 SQGFGAATNSGPLELSTVSGSYWAGVPLYTSKE-----KSTP-INMTFVDSMDAYL 388
DB 350 SQTGFSAHYSQDVMVGAVDWNGTVVWQKASQIIIPRNTFNVSTKKNPL-ASTL 408
QY 389 GYAAAIILNRVQSL-VLGAAPYQHIGLVAMPQNTGMNESNANVGTQIGAYFGASLGS 447
DB 409 GYTVNSATASSGDVLYIACOPRYNHTGVIIYRMEDGNKILQTLSEGLSGSYLSIT 468
QY 448 VDVSNGSTDLVLIGAPHY-----YEQTR-GGQVSVCPPLPRGORARWQCD 491
DB 469 TDIDKDSNTDILLVGAFWYMGTEKEBQGVVYVALNQTRFEYQWLEPIKQTCSSROH 528
QY 492 AVLYGSGQGPWG-RFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVYLFHGTSGSGISP 550
DB 529 SCTTENKNEPCGARFGTAIAAVKDLNLDGPNDIVIGAPLEDGGAVYIYG-SGKTIK 587
QY 551 SHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRLSQPVLVRKAIM 610
DB 588 EYAQRIPSGDGTLLKFPFGQSIHGEDMLNGDLTDTVTIGLGAALFWSRDVAVVKVTWN 647
QY 611 PNFREVARNVFECNDQVVKKEAG--EVRVCLHVO-KSTEDRLREGQIQSWVTVDLALDS 667
DB 648 FEPKNVNIQKNCH--MEGETVCINATVCFEVLKSKEDTIYEADLQ-----YRVILDS 700
QY 668 GRPHSRVAFNET-----KNSTRQTOVLGELTQCTELKLQLPNCIEDPVSPIVRLNFS 721
DB 701 LRQISRSFFSGTQERKQVNRITVRSKSEC-----TKHSFYMLDKHDPODSVR---ITLDFN 752
QY 722 LVGTPLSAFNLRPVLADRAQLFTALFPEKQKCGNDNIQDDLSITFSPMSLDCLVVG 781
DB 753 LT-DPENG-----PVLDDSLPNSVHEYIPFAKQCGNKKKICISLSLHVATTEKDLIVRS 806
QY 782 PRE-FNVTVVRNDGESSYQTVTFPPPLDLSYRKVSTLQNRQSQRSLRACESSSTEV 840
DB 807 QNDKFNLSLTVKNTKDSAYNTRIVHYSPLNVSFGIEAIQKD-----SCESN----- 853
QY 841 SGALKSTSCSINHPIFENSEVFNITTFDVSQASIGN-KLLKANVTSENMPRTAKTE 899

RESULT 8

I45914
Integrin alpha 2 subunit - bovine (fragment)
C:Species: Bos primigenius taurus (cattle)
C:Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 15-Sep-2003
C:Accession: I45914
R:Kanata, T.; Puzon, W.; Takada, Y.
J. Biol. Chem. 269, 9659-9663, 1994
A:Title: Identification of putative ligand binding sites within the I-domain of integrin
A:Reference number: A54402; MUID:94193647; PMID:7511592
A:Accession: I45914
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1170 <KAM>
A:Cross-references: GB:I45886; NID:9439695; PIDN:AAB59255.1; PID:9439696
P:161-336/Domain: von Willebrand factor type A repeat homology <VWA2>

Query Match 18.2%; Score 1071; DB 2; Length 1170;
Best Local Similarity 27.4%; Pred. No. 2.4e-66;
Matches 333; Conservative 219; Mismatches 495; Indels 168; Gaps 47;

QY 1 FNLDTENAMTFQ-ENARGFGQSVVQL---QGSRRVVVGAPOEIVAAQNRGLSQCDY 54
DB 19 YNVGLPKAKIFGSPSSQFGYAVQQFINPKGNWLLVGSFWSGFPKRRMGDVYKCPVDLST 78
QY 55 GSCEPIRLQ-----VPVEAVNMSLGLSLAATTSFPQLACGPTVHGTCSNTYVKG 107
DB 79 TTCXKLNLTQSTMSNVTETKTNKMSLGLTLTRNVGTGGFLATCCPLMAQCCGQYVYTGVC 138
QY 108 FLFGNLRQOPQOKPPEALRGCPQEDSDIAFLIDGSGSIIPDFRMEKWSVTWVQ 165
DB 139 SDVSPDP-QLRTSPAPAVQTCPSFIDVVVVCDESNISYPWD--AVKPFLEKXVQGLDIG 194
QY 166 KSKTLFSLMYSSEFRHFTFKFQNNPNPSRLIKPTQLL---GRTHATGLRKVRNE 221
DB 195 PTKTQMGLLIQANNRVVFNLTNFKSKD---EMIKATSTQTPQYGGDLTNTPKAIQVARDT 251
QY 222 LFNITNGARKNAFKILFLITDGEKFGDPLGYEDVPELDREGVIRYVIGV-----GDAFR 276
DB 252 AYSTAAGRPGCATKVVVVVVDGESH-DGSKLKAVIDQCNKDNILRFGIAVLGYLNRNALD 310
QY 277 SEKSRQELMTVASKPRPDHVFQINNFEALKTIONQREKIPAIPTGTQSGSSPFHEMSQ 336
DB 311 TKNLKEIKALIASITERHFPNVSEADLLEKAGTIGQIFSIETVQGGDNFQWMSQ 369
QY 337 EGFSAAIT--SNGPLLSTVSGYDAGGVFLYTSKEKSTFINMT--RVSDSMN-DAYLGYA 391
DB 370 VGFSAEYSQNNILMUGAVYDWSGTVVQKTPHGLIFSKOAFEQILQDRNHSYLYGS 429

[illegible]

RESULT 9
S44142
VLA-2 protein homolog - mouse
C:Species: Mus musculus (house mouse)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 15-Sep-2003
C:Accession: S44142
R:Edelman, J.M.; Chan, B.M.; Uniyal, S.; Onodera, H.; Wang, D.Z.; Damjanovich, L.; Latze
submitted to the EMBL Data Library, January 1994
A:Description: The mouse VLA-2 homologue supports collagen and laminin adhesion but not
A:Reference number: S44142
A:Accession: S44142
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1178 <DE>
A:Cross-references: EMBL:Z29987; NID:9473098; PID:CAAB2877.1; PID:9473099
P:169-344/Domain: von Willebrand factor type A repeat homology <VWA2>

```

Query Match      18.0%; Score 1057; DB 2; Length 1178;
Best local Similarity 27.8%; Pred.No.2.3e-65;
Matches 340; Conservative 210; Mismatches 487; Indels 188; Gaps 44;

QY      1 FNLDTENAMTFQ--ENARGFQGSVVQL----QGSRVVVVGAPQETIVAAQNQRSLYQC--DYST 54
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      27 YNVLPGAKIPSGSSBOFGYSVQQLTNPCQNWLLVGSPPWGPENRMGDVYKCPVDLPT 86

QY      55 GSCPEIRLO-----VPVEAVNMSGLSLAATTSPPQLLAGCPTVHQTCSENTVYKGLC 107
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      87 ATCEKLNLNQNASISNVTETKTNMSLGLLITRNPGTGGFLTCGLWAHQCNQYATGIC 146

QY     108 FLFGSNLFEQQPQ---KEPPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEQL 164
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB     147 ---SDVSPDPQFLTSPSAVQACPSL--VDVVVVCDESNSIYP--WEAVKNFLAVFVTVGL 199

QY     165 K--KSKLPLSLMOYSESEFRTHFTFKFQNNPNPSRLIKPIITOLLG--RTHVATGLAKVKRE 221
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB     200 DIGPKKTKUVALIQVIANPFIIFNLNDFETKEDVMQATSETRQHGDLTNVFAIFPARDY 259

QY     222 LFNITNGARKNAKIFLLITDGEKPGDPLGYEVDIPELDRBGVIRVIGV-----GDAPR 276
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB     260 AYSQTSQGRPGATKVMVMVVDGESH--DGSCLKTVIQCCNDDEILRFGLIAGLYLARNALD 318

QY     277 SEKSQELNIVASKPPRDHFVQINNPEALKTIQNQLREKIFAEGTQTCSSSPFEHMQSQ 336
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB     319 TKNLKEIKALASTPTERYFNVADEALLEKAGTIGEQIPSIIEGTIVQG--GNFQMEKVAQ 377

QY     337 EGESA--AITGSPLLSTVSGSYDWAGGVFLYTSKEKSTFINMT--RVDSDMN--DAYLGVA 391
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB     378 VGFSADYAPQNDILMLGAVGAPDWSGTVLQETSHKPIPPKQAFDQVLQDRNHSSFLGYS 437

QY     392 AAILLENRVQSILVIGAPRYQHIGLGVAMFRONTGMWESNANV----KGTQIGAYFGASLCS 447
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB     438 VAAIETGCVHPVAGAPRANTTQGIULIYSNK---QGVTVIQSHRGDQLGSGYEGSVLCS 494

QY     448 VDVDNNGSTDLVLIGAPHYEYBQTR--GGQVSVCLPFGQBARQCDAVLGEGQQGWGRP 505
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB     495 VDVDKDTITDVLVAGAPTYMDLKBBEGKYVLFITIKGLNHQ---FLEGPEGTQGNAPF 551

QY     506 GHALTVLDGVNDKLTDLVAIGAPCEENRGAVLYFHTGTSGSGISPHSQRIAGSKLSPR--564
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB     552 GSAIAALSDDNDGFDVIVGSPVENENSGNVIIYNGHQT--ITKYSQKILGNSGAFFR 610

QY     565 -LQYFQSGISGGQDITMDGLVDLTVGAQGHVLLRLRSQPLVRVKAIMEFNPREVARVPEC 623
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB     611 HLQFGRSLDGYDLNGDSITDVSIGALGVQIQLWSQSIAADVAIEALFTP-----660

QY     624 NQGVKGEAGVRVCLHVOKSTDRLRREGQIQSWVTVYDLALD----SGRPHSRAVFNET 679
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB     661 -DKITLLNKDAKITKLCFRAEFPAGANNQV--AILPNMLDADGHSSRVTSGVGFREN 717

QY     680 KNSRTRQTVGLGTVQCFET--LKQLPNCIEDPVSPIVLRINPSLVCTPLSAFNLRPVL 737
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB     718 SERFLQKNVMVNEVQCKSEHHISIQKPS--DVVNPLDLAVDISLENPGTS-----PAL 768

QY     738 ABDAQRFLTALPPEKNGCNDNIQODDLISI-----TFSPMSLDCLVGVGPFENVT 788
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB     769 EAYSEVTKVPSLPFYKEGSDGICISDLILDVQQLPALQTSF-----IVSNQKMLTFS 823

QY     789 VTVRPDGRDSYQTVTPFPFPLDLSYRKVSTIQNORSORSRWELACESASST-EVSGALKST 847
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB     824 VILKRGESAYNTVVLAFSENLFP-----ASFSPMDVGTTEVTCVGGSQKSV 871

QY     848 SCGINPIPFENSEVTFNTTPOVDSKASLGKLLKANVTSENNMPTNKTE--FOLELP 905
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB     872 TCDVGYPALKSEQQVTFINFDNLQ--NLQNAQAINQAFSESQ--ETNKADNSVSLTIP 928

QY     906 VKYAVTVMTVSHGVSTKYLNFTASNTSERVMQHOYQVSNLQOR-----SLPSLV 955
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB     929 LLYDAEHLHT--RSTNINFTYISDDNAPSATVK---SVEDGPKFIFSLKACGAPVSM 984

```

496 IDIDKDSYTDLLLVGAPMVGTEKEQGKYVAVNQTRPEYQMSLEPIRQTCCSSLKDN 555
 QY AVLYGEQQPGWG-RFGAALTVLGDVNGDKLTDVAIGAPCEEDNRGAVYLPHGTSGSGLSP 556
 :
 556 SCTKENKEPFCGARFETAATAVKDLNVDGFNDVWICAPLEDHAGAVIYHG-SKKTIRE 614
 DB :
 QY SSSORTAGSKLSPRLGYFGOSLSGGDILMDGLVDLTGCAQGHVTLTLLRSQPVLVRKAIME 610
 :
 615 AVAQRIPLSGDGDKTYLKFFQSISGENDLNGDLTVTTIGLGGAALFWARDVAVVKVTMN 674
 DB :
 QY FNPFRVARNVFECDNDGVKKGBAG--EIVAVCLHVQ-KSTRDRLRREGOIQSVVYDIADLS 667
 :
 675 FBPNKVNIQKGNCR--VECKETVCINATMCTHFVULKSKEDSIYEADLQ--YRVTLDS 727
 DB :
 QY GRPHSRVAFNET-----KNSTRTTOTVGLTGTCCTKLQLPNCI-----EDPVS 712
 :
 728 LQGISSEFPSTGQERKIQRNITVRSE-----CIRHSFYMLKHDFQD 770
 DB :
 QY PIVLRNLFLSVGTPLSAFNLRAPLVALABDAORLFTALFPPEKCNCGNDNI CODDLSITFSFM 772
 :
 771 SVRYTLDNFLT-DPENGL-----PVLD DALPNSVHEHIIPFAKCGNKERICISLTLNVSTT 834
 DB :
 QY SLDCILVVGGPRE-FNVTVTRVNDGEDSYRTQVTFPPPLDLSYRKVSTLQNQRSQRWRLA 831
 :
 825 EKSLLIYRSQHDKFNWSLTVKKKGDSAYNTRIVVOHSPNLIIFSGIEIQKD-----S 876
 DB :
 QY CESASSTEVGALKSTSINSIHPIIPENSEVITFNITFDVDSKASLGKLL-LKANVTSEN 890
 :
 877 CESN-----QNITCVGYPFIRAGETVTFKIIPQNTSHLSENAIHLISATSDSEE 927
 DB :
 QY NMPRTNKTFOELPCKYAV---YNVVTSHGVST-----KYLNFTASENSTRVMOHQ 939
 :
 928 PLESINDNEVNISI PVKEVGLQFYSSASEHHISVAANETIPEFINST-EDIGNEINVF 985
 DB :
 QY YOVSNLQORSLP---ISLVP-----IYVRLNQTVMID-----RP----- 971
 :
 986 YTIKRKGHPMPBELQSLTSFPMLTAGYVLYPIG-----WSSSDNNVRCPRSEDPPG 1039
 DB :
 QY -----QVTPS-----ENLSSTCHTKERLPSHSDFLABELKAPVNVCSIAVCORIQ 1016
 :
 1040 INSGKMTIISKSEVLKRGHTIQDCSTC-----GVATITCSLLPSDLSQ 1082
 DB :
 QY CDL-----PPGIQEEP---NATLKNLSPDWYIKTSHNHILLIYSTABILFNDYSVFTLL 1067
 :
 1083 VAVSLLLWKPTF-IYAHFSSLNLTJRGELX-----SENSLTLSSSN----- 1123
 DB :
 QY PQCGAFVRSQOTKVEPPEVNPPLPE--IVGSVSGGLLALLITALYKLGFFKRYQKDM 1125
 :
 1124 -----RKEELAIQISKDGLPCRVELWTILSAPAGLLALLMLLAILAKWIQGFPRPLKKK 1177
 DB :
 QY 1126 MSE 1128
 | :
 DB 1178 MEK 1180

RESULT 11
 A33998
 N:Alternates alpha-2 chain precursor - human
 N:Alternate names: CD49b; platelet glycoprotein GPIa; VLA-2/collagen receptor a
 C:Species: Homo sapiens (man)
 C>Date: 30-Mar-1990 #sequence revision 18-Sep-1992 #text_change 15-Sep-2003
 C:Accession: A33998; B56793; A53117
 R:Takada, Y.; Hemier, M.E.
 J. Cell Biol. 109, 397-407, 1989
 A:Title: The primary structure of the VLA-2/collagen receptor alpha(2)-subunit
 A:Reference number: A33998; PMID:89308879; EMBL:2545729
 A:Accession: A33998
 A:Molecule type: mRNA
 A:Residues: 1-1181 <Trk>
 A:Cross-references: GB:I17033; NID:g33906; PID:CAA34894.1; PID:g33907
 A>Note: the authors translated the codon GAT for residue 802 as Gln, GTC for res
 B:Catimel, B.; Parmentier, S.; Leung, T.L.; McGregor, J.L.
 Biochem. J. 279, 419-425, 1991


```
Db 218 VMGAPGSSWTGSLFYNTTKYKFLDKQVQKF---GSLYGVSGAGHFRSQTHTTV 274
Qy 404 VLGPARYQHIGLVAMFRONTGMWESNANV---KGTQIGAYFCASLCSVDVDSNSTDI 459
Db 275 VVGAPQHEGIGKAYIF---SIDEKELTLHENKGLKGLSYFCASVCAVDLNAQFSDL- 329
Qy 460 LIGAPHYVETRGGOVSVCPLPGQARWC-DAILYGEGQGPWGPFGAALTVLGDVNGD 518
Db 330 LVGAPQWSTIREGRVYV-INGSGAVNMMETNLVGSCKYA-AHFGEISVNLGDIDND 387
Qy 519 KLTVDVAGPGEEDNRGAVYLFHGTSGSISPHSQRISAGSKLSPRLQYFGQSLGGQDL 578
Db 388 GPEDVAIGAPQEDDLGAIYVNGRA-DGSSFTFSORIBGLQISKLSMFGSISQIDA 446
Qy 579 TMDCLVDLTGCA--QGHVLLRSQPLRVKAMFNPREVARNVPCNDQVKGKAGEV 636
Db 447 DNNGYVDVAVGAFRSDSAVLLRTPVIVDASLS-HPESVNRKTKFC-----VENGP 498
Qy 637 RVCLHVOKSTRDRRLREGQIOSVVYDLALDSGR-----PHSRAVENETKSTRRTQVLG 691
Db 499 SVCIDITLCSYKKEVPGYIVLFYKMSLDVNRKABSPRFYSSNGTSDVITGSIQVS 558
Qy 692 LFTCTETLKLQPNCTEDVPSIVLRLNLSLVTPLS-----AFGNLRPVLAEDRQ-LF 745
Db 559 REANCRTHOAFMEKVDRIILPTQIEAAVHLGPHVTSKSTEEFPPLQILOQKKEKDIM 618
Qy 746 TALFPPEKQGNINICODDLSIT--PSFM-----SLDCLVVGPREPNTVTVRNDGDSY 799
Db 619 KKTINARFCAHEN-CSALQVSAKGLPKPHENKTYLAVGSKMTLNLVSLFNAGDDAY 677
Qy 800 RTQVTFPPDLISYRVKSTLQNSORSWRMLACESASSTEVSGALKSTSCSINHPIFPPN 859
Db 678 ETLHLVKLPGLYFIKILEEK-----QINCE---VTDSNGVQ-LDCSIGIYVDHL 727
Qy 860 SEVTFNITPDVDSKASLGNKLLKANVTSEN-NMPTKTKTPEQLEPVKYAVYVWVTS 917
Db 728 SRIDISFLDVSLSRABEDLSITVATCENBEENDNLKHSRVTVAPLKYEVKLVHGF 787
Qy 918 GVSTKYLNTASNTSRV-----MCHQYQVSNLQORSLP-ISLVFLVPVRLNQTVIMDRP 971
Db 788 VNPITSFVYSGNDENEPETQVMEKMLTTHVINTGSMAPNVSVIIMVPSFS-----P 840
Qy 972 QVTFSENL-----STCHTKERLPHSDPLAELRKAPVNVCSIAVCQRIQDIPFGIQ 1025
Db 841 QTDKLFNILDVQTTGCEHFNYQ-----RVCALQ-----Q 872
Qy 1026 EEFNATLKNLSFDWIKTSHNLLIVSTAS-----ILFN-----DSVFTLLPGQG 1071
Db 873 KSMQTLKGVRF---LSKTDKELLICYADPHCLNPLCNFGKMSGKESVHIQLEGEP 929
Qy 1072 AFVRSQTEKVPFEV-----PNP-----LPLIV 1095
Db 930 S-ILEMDETSAKFEIRATGFPENPVRVIELNDENVAHVLLGLEHHPKRYFTVILIS 988
Qy 1096 GSSVGGHLLALITAAKYKLGKPKQYKOWMSR 1128
Db 989 SSLLLGLVILLISYVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWV 1021

RESULT 14
158409
integrin alpha-9 chain precursor - human
C:Species: Homo sapiens (man)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 31-Dec-2000
C:Accession: 158409; A49459
R:Hibi, K.; Yamakawa, K.; Ueda, R.; Horio, Y.; Murata, Y.; Tamari, M.; Uchida, K.; Takah
Oncogene 9, 611-619, 1994
A:Title: Aberrant upregulation of a novel integrin alpha subunit gene at 3p21.3 in small
A:Reference number: 158409; MUID:94119603; PMID:8290272
A:Accession: 158409
A>Status: preliminary; translated from GB/EMBL/DBSJ
A:Molecule type: mRNA
```

```
A:Residues: 1-1035 <RES>
A:Cross-references: GB:D5303; NID:9464180; PIDN:BAA04984.1; PID:9533327
R:Palmer, E.L.; Ruegg, C.; Ferrando, R.; Pytela, R.; Sheppard, D.
J. Cell Biol. 123, 1289-1297, 1993
A:Title: Sequence and tissue distribution of the integrin alpha9 subunit, a novel partn
A:Reference number: A49459; MUID:94064789; PMID:8245132
A:Accession: A49459
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 30-1035 <PAL>
A:Cross-references: GB:L24158
C:Superfamily: integrin alpha-4 chain
C:Keywords: glycoprotein; metal binding; transmembrane protein
F:1-27/Domain: signal sequence #status predicted <SIG>
```

Query Match 10.7%; Score 630; DB 2; Length 1035;

Best Local Similarity 26.5%; Pred. No. 1.6e-35;

Matches 237; Conservative 160; Mismatches 332; Indels 164; Gaps 40;

```
Qy 331 EHMSQEGPSAAITNGNGLLSTVGSYDWAGOVFLYTSKEKSTFFINMT-RVDSMDNDAYLG 389
Db 190 EHSGCOAGIAGFFTEELVVMGAPGSPYWGATIKVNLTD-NYLLKLNDEVIMNRYTYLG 248
Qy 390 YA-AAITLNRVQSLVLGAPRYQHIGLVAMFR--QNTGMWESNANVGTQIGAYFGASL 445
Db 249 YAVTAGHSHPSITIDVVGAPQDKGIGKGYITFRADRSGLTIKIFQASGKMGSYFGSSL 308
Qy 446 CSVDVDSNGSTDLVLIGAPHYVETRGGOVSVCPLPGQARWCDAVLYGEGQGPWGRF 505
Db 309 CAVDLNEDGLSDL-LVGAPMFSEIRDEQVTVY-INGNGALEE-QLALTGD-GAYNAHF 364
Qy 506 GAALTVDVNGDKLTDVAIGAPGEEDNRGAVYLFHGTSGSISPHSQRISAGSKLSRPL 565
Db 365 GESTASLDLNDGFPDVAIGAPKEDDPAGAVYIYHGDAG-GIVPQYMSLGGQKINPVL 423
Qy 566 QYFGSLSGGODLTMDGLVDLTGCA--QGHVLLRSQPLRVKAMFNPREVARNVPEC 623
Db 424 RMFGISGGIDMGNGYDVPVVGAFMDSVLLRARPVITVD-VSIFLPSINTAQC 482
Qy 624 NDQVVKGEAGEVRV-CLHVQ-----KSTRDLREGQIOSVVYTDIALDSGRPHSRAV 675
Db 483 HD-----QQPWNCLNVITCFSGHKHVPBEI---GLNVLMADVAKKKGQMPRVY 531
Qy 676 F---NETKNSTRTOVLGLTQCTETLKLQPNCTEDVPSIVLRLNLSL-----VGTPLS 728
Db 532 FVLLGETMGQVTKLQTYMEETCRHYVAHVKKRVQDVISPIVFEAATSLSEHVTVGSEER 591
Qy 729 AFGNLRPVL-----AEDAQRLLFTALFPPEKQGNINICODDLSITFSFM--SLD----CL 777
Db 592 ELPLPLTVLRWKKKGKIAQKNQTV---PERNCSED-CAADLQLOGKLLSSMDKLTLYL 647
Qy 778 VVGGRPRENVTVRNDGDSYRTQVTFPPDLISYRVKSTLQNSORSWRMLACESASS 837
Db 648 ALGAVKNSILNISINLGLDDAYDANVSNVRELFF-----INMQKEEMGISCELLAES 701
Qy 838 TEVSGALKSTSCSINHPIFPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPNPK 897
Db 702 DFL-----KCSVGFPPARSKSKYFSVIFDTHSL--GEDEVLSFVITAQSG--NTER 750
Qy 898 TS-----FQLELPVKYAVYVWVTS-----HGVSTKYLNTASNTS---RVMQHOY 940
Db 751 SESLHDNTLWMLVPLMHEVDTSITGIMSPTSFFVYGESVDAANFIQDLDECHQFQINIL 810
Qy 941 QVSNLQORSLPISLVFL-VPVRLN-----QTVIMDRPQVTFSENLSSTCHTKER 988
Db 811 QVYNTGPSTLPGSSVSISFPNRLSGGAEMFHVQEMVVGQKNGCSFQKNTPFCIIPQEQ 870
Qy 989 LPSHSDFLAELRKA-----PVNVCIAVCQRIQDIPFGIQEENFATLKNLSRPL 1039
Db 871 ENIPTFTAFPTKSGRKYDCKEKGISCLTAHCN-----FSALAKESNTI-----D 917
Qy 1040 WTIKTSHNLLIVSTABILFNDSVFTLLPGQAFVRSQTEKVP----- 1084
```

Db 918 1Y-----MLANT-BILKDDSSVIO-----FMSRAKVKYDPALRAVVEIAHGNPEEV 962
Qy 1085 ---FEVPR---PLPLIVG-----SSVGGLLLLALITAAALYKLGFFKQYKXMM 1126
Db 963 TVTFALENLPRGVVGVGWIATISLLVGLIFLLAVLLVWNGFFRRRYKEII 1015

RESULT 15
T31437
integrin alpha chain SU2 - sea urchin (Lytechinus variegatus)
C:Species: Lytechinus variegatus (variegated urchin)
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jan-2000
C/Accession: T31437
R:Hertzler, P.L.; McClay, D.R.
submitted to the EMBL Data Library, May 1998
A:Description: Alpha SU2, a sea urchin integrin which binds laminin.
A:Reference number: Z21035
A:Accession: T31437
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1041 <HER>
A:Cross-references: EMBL:AF067658; NID:g3220240; PID:g3220241; PIDN:AAC23572.1
A:Experimental source: developmental stage embryo
C:Function:
A:Description: binds laminin
C:Superfamily: integrin alpha-2b chain

Query Match 10.4%; Score 614.5; DB 2; Length 1041;
Best Local Similarity 24.8%; Pred. No. 2e-34;
Matches 236; Conservative 155; Mismatches 314; Indels 245; Gaps 41;

Qy 336 QEGFSAALITNGP--LLSTGVSYDWAGGVFLYTSKEKSTFINTVRVDS-----DMNDAYLG 389
Db 181 QAGSGIIFSDNSALVWAPGSYVLOGQIYVQSLLNRV-VQATQESNTGYTFSDNSYRG 239
Qy 390 YAAAI--ILNRVQSLVIGAPRYOHI-GLVAMFRONTGMNESNANVKTQIGAYFGASLC 446
Db 240 YSLALGDFNGGVQDYYVGTTPRAESLMGLVAIFDQNLNQN-----QVMGTQIVAYFGYSVT 296
Qy 447 SVDVDSNGSTDLVLIGAPHYEQTEGGGVSVCLPRGQARMCQDVALYGEQ----- 498
Db 297 VVDI--NNDYDOLLVGAAPWMDGPAIQ-----RWEAGAVVYVQLNPDVGPGA 343
Qy 499 -----GQWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLFHGTSGS 546
Db 344 SNRLSLSTLIGGQIRSRFGLSIASIGDSNQDGFNDVAIGAPYEGDDAGAVYVHG-SAN 402
Qy 547 GISPSHSQRIAGSKLS-PRLOVFGQSLSCGQDLTMDGLVDTVGAQ--CHVLLILRSQVVL 603
Db 403 GLKSTPAQVLTPSTLGHSGITTFGFSLQGGQMDKNKYFDLLVGAESANTAVLIRTPWV 462
Qy 604 RVKALMEFNPREVARNVPECNDQVYKGEAGEVRVCLHVOKSTRDRLRREGQIQSV----- 658
Db 463 SLDATLNTPE-----IGINLENKTYE-LADGTMTVTSFIAMT 497
Qy 659 -----VTYDLALDSG-RPHSRVFNTRKNSRROTQVTLGL-TQTCTELKQ 702
Db 498 CFTYTGNYLPDHIDISYTVTVDSGLIANRRAMFVNDMSSEITKTRRLAVSTQFCDPLEAY 557
Qy 703 LPNCIEDPVSPTVLRNLNLSVCTPLSAFGN-----LRPLVAEDAQRLLTALFPPEK 753
Db 558 VGNSTEDKLTFLKVLQYDL-----NDESRLOPHEIIPIDMATMSTQTKQVSIQN 609
Qy 754 NCGNDNICQDLSITFSFMSLDCLVVGPREFNVTVVRNDGEDSYRTQVTFPPPLDLSY 813
Db 610 NCVN-NICIPDLDTVT-PNLPNIVIGQTQELTLDVSLNNGEDAFQSSLSVYVPLGLQF 667
Qy 814 KXVSTLQNRQSRWELACASASSTEVSGALKSTSCSINHPIFEN-----SEVTENIT 867
Db 668 VRL-----ERKANMDSFVTCSDSD-----LRITCDTGNPMVGKNILBFGILTSTFQVS 717
Qy 868 FDVDS-----KASLGNKLLKANVTSENMPRTNKTEFQLELPVKYAVYVMTVSHGVSTK 922

Db 718 GKDSIEFYFKAESSENS--EDPNTLNNELNMTVPVTVDCITLKLKLSASYEIVMYSTQED 775
Qy 923 YL-----NFTASENTSRVWQHQYQV-----SNLGQSLPIS-----LVFLVPV 960
Db 776 YVVPFPKNAASEADIGMEVM-HLYEVRNTGSSNAGEVSLNIQWPKQNEGEVLYLLGI 834
Qy 961 RLNQTVINDRPQ-----VTFSENLSTCHTKERLPSHSDFLAELRKAPVV 1005
Db 835 MTEBEGVTCLTOGKANPEGVKLEPSTKAKLSNSTTQVSGRKRREPEVAEALAQTDN--VI 892
Qy 1006 NCSIAVCORIQCQDIPFFGIOBEFNAT-----LKGMLSPDWYIKTSHNHLILVSTAILF 1059
Db 893 YCASDSQVLIINCTI-----DEINASKSVVRLGRF--W-----ERTF 928
Qy 1060 NDSVFTLLPQGGAPVRSSQTETKVE--PFEVNP-----LP----- 1092
Db 929 QKAVSELTPVVOATIASASAARKTIPYINPLPRDFSDSTKASTLVTEELVPPVTPIAW 988
Qy 1093 -LIVGSSVGGLLLLALITAAALYKLGFFKR-----QYKDMMS--EGGPP 1132
Db 989 WIIVWSVLGGIILLILLGLMKCGFFERKKPGEERKAYAVASADKGGPP 1038

Search completed: June 7, 2004, 17:18:02
Job time : 38.559 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 16:55:38 ; Search time 9.97194 Seconds
(without alignments)
5937.039 Million cell updates/sec

Title: US-09-902-481B-4

Perfect score: 5884

Sequence: 1 FNLDTENATFQENARGFGQ.....FKRQYKDNMSRGPPGAPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5836.5	99.2	1152	1 ITAM HUMAN	P11215 homo sapien
2	4460	75.8	1153	1 ITAM MOUSE	P05555 mus musculus
3	3459	58.8	1163	1 ITAX HUMAN	P20702 homo sapien
4	3401	57.8	1162	1 ITAD HUMAN	Q13349 homo sapien
5	1547.5	26.3	1170	1 ITAL HUMAN	P20701 homo sapien
6	1527.5	26.0	1163	1 ITAL MOUSE	P24063 mus musculus
7	1148.5	19.5	1167	1 ITAE MOUSE	Q60677 mus musculus
8	1140	19.4	1179	1 ITAE HUMAN	P38570 homo sapien
9	1093.5	18.6	1151	1 ITAL HUMAN	P56199 homo sapien
10	1084.5	18.4	1189	1 ITAH HUMAN	Q9UKX5 homo sapien
11	1071	18.2	1170	1 ITAH BOVIN	P53710 bos taurus
12	1057	18.0	1178	1 ITA2 MOUSE	Q62469 mus musculus
13	1055	17.9	1180	1 ITA1 RAT	P18614 rattus norv
14	1054	17.9	1181	1 ITA2 HUMAN	P17301 homo sapien
15	1051.5	17.9	1167	1 ITAG HUMAN	Q75578 homo sapien
16	663	11.3	1039	1 ITA4 MOUSE	Q00651 mus musculus
17	642	10.9	1038	1 ITA4 HUMAN	P13512 homo sapien
18	630	10.7	1035	1 ITA4 HUMAN	Q13797 homo sapien
19	593.5	10.1	1032	1 ITA4 XENLA	Q91687 xenopus lae
20	571.5	9.7	1066	1 ITA3 CRISP	P17852 cricetidae
21	567.5	9.6	1053	1 ITA3 MOUSE	Q62470 mus musculus
22	555.5	9.4	1053	1 ITA5 MOUSE	P11688 mus musculus
23	546.5	9.3	1034	1 ITAV CHICK	P26008 gallus gall
24	545.5	9.3	1050	1 ITA5 XENLA	Q06274 xenopus lae
25	537.5	9.1	1130	1 ITA6 HUMAN	P23229 homo sapien
26	535	9.1	1044	1 ITAV MOUSE	P43406 mus musculus
27	534.5	9.1	1072	1 ITA6 CHICK	P26007 gallus gall
28	532.5	9.0	1066	1 ITA3 HUMAN	P26006 homo sapien
29	532	9.0	1049	1 ITA5 HUMAN	P08648 homo sapien
30	526	8.9	1048	1 ITAV HUMAN	P06756 homo sapien
31	517	8.8	1044	1 ITA8 CHICK	P26009 gallus gall
32	512.5	8.7	1091	1 ITA6 MOUSE	P61739 mus musculus
33	498	8.5	1179	1 ITA7 MOUSE	Q61738 mus musculus

ALIGNMENTS

RESULT 1

ID	ITAM HUMAN	STANDARD	PRT	1152 AA
AC	P11215			
DT	01-JUL-1989 (Rel. 11, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor Mo1)			
DE	(Neutrophil adherence receptor)			
GN	ITGAM OR CR3A OR CD11B			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RX	SEQUENCE FROM N.A.			
RX	MEDLINE=883115033; PubMed=2457584;			
RA	Corbi A.L., Kishimoto T.K., Miller L.J., Springer T.A.;			
RT	"The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CD11b) alpha subunit. Cloning, primary structure, and relation to the integrins, von Willebrand factor and factor B.";			
RL	J. Biol. Chem. 263:12403-12411(1988).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RP	MEDLINE=88190151; PubMed=2833753;			
RA	Arnaout M.A., Renold-O'Donnell E., Pierce M.W., Harris P., Tenen D.G.;			
RT	"Molecular cloning of the alpha subunit of human and guinea pig leukocyte adhesion glycoprotein Mo1: chromosomal localization and homology to the alpha subunits of integrins.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 85:2776-2780(1988).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88257215; PubMed=2454931;			
RA	Arnaout M.A., Gupta S.K., Pierce M.W., Tenen D.G.;			
RT	"Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor Mo1 (complement receptor type 3).";			
RL	J. Cell Biol. 106:2153-2158(1988).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=93123748; PubMed=8419480;			
RA	Fleming J.C., Pahl H.L., Gonzalez D.A., Smith T.P., Tenen D.G.;			
RT	"Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-integrin gene family demonstrate remarkable conservation of genomic organization and suggest early diversification during evolution.";			
RL	J. Immunol. 150:480-490(1993).			
RN	[5]			
RP	SEQUENCE OF 9-1153 FROM N.A.			
RX	MEDLINE=89098993; PubMed=2563162;			
RA	Hickstein D.D., Hickey M.J., Ozols J., Baker D.M., Back A.L., Roth G.J.;			
RT	"cDNA sequence for the alpha M subunit of the human neutrophil adherence receptor indicates homology to integrin alpha subunits.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 86:257-261(1989).			
RN	[6]			

34	494	8.4	1396	1 ITA2 DROME	P12080 drosophila
35	491.5	8.4	1146	1 ITAL DROME	Q24247 drosophila
36	490	8.3	1033	1 ITAB MOUSE	Q9QUM0 mus musculus
37	489.5	8.3	1039	1 ITAB HUMAN	P08514 homo sapien
38	489	8.3	1025	1 ITAB HUMAN	P53708 homo sapien
39	486	8.3	126	1 ITAM CAVEO	P11578 cavia porce
40	471	8.0	1181	1 ITA7 HUMAN	Q13683 homo sapien
41	470	8.0	1106	1 ITA7 RAT	Q63258 rattus norv
42	462	7.9	1226	1 PAT2 CAEEL	P34446 caenorhabdi
43	444.5	7.6	1139	1 INAL CAEEL	Q03600 caenorhabdi
44	427	7.3	1115	1 ITA3 DROME	Q43386 drosophila
45	385	6.5	1000	1 ITA5 DROME	Q9W1M8 drosophila

RP RX MEDLINE=92073318; PubMed=1683702;
RA Shelley C.S., Arnaut M.A.;
RT "The promoter of the CD11b gene directs myeloid-specific and
RT developmentally regulated expression";
RL Proc. Natl. Acad. Sci. U.S.A. 88:10525-10529(1991).
RN [7]
RP RX SEQUENCE OF 1-9 FROM N.A.
RP RX TISSUE=Blood;
RP RX MEDLINE=92144986; PubMed=1346576;
RA Pahl H.L., Rosmarin A.G., Tenen D.G.;
RT "Characterization of the myeloid-specific CD11b promoter.";
RL Blood 79:865-870(1992).
RN [8]
RP RX SEQUENCE OF 17-31.
RP RX MEDLINE=87076671; PubMed=3539202;
RA Pierce M.W., Remold-O'Donnell E., Todd R.P. III, Arnaut M.A.;
RT "N-terminal sequence of human leukocyte glycoprotein Mol.
RT conservation across species and homology to platelet IIb/IIIa";
RL Biochim. Biophys. Acta 874:368-371(1986).
RN [9]
RP RX X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 148-331.
RP RX MEDLINE=95171458; PubMed=7867070;
RA Lee J.O., Rieu P., Arnaut M.A., Liddington R.;
RT "Crystal structure of the A domain from the alpha subunit of integrin
RT CR3 (CD11b/CD18).";
RL Cell 80:631-638(1995).
RN [10]
RP RX X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 148-334.
RP RX MEDLINE=96363671; PubMed=8747460;
RA Lee J.O., Bankston L.A., Arnaut M.A., Liddington R.C.;
RT "Two conformations of the integrin A-domain (I-domain): a pathway for
RT activation?";
RL Structure 3:1333-1340(1995).
RN [11]
RP RX X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 148-337.
RP RX MEDLINE=98362595; PubMed=9687375;
RA Baldwin E.T., Sarver R.W., Bryant G.L. Jr., Curry K.A.,
RA Fairbanks M.B., Finzel B.C., Garlick R.L., Heinrikson R.L.,
RA Horton N.C., Kelley L.L., Mildner A.M., Moon J.B., Mott J.E.,
RA Muchler V.I., Tomach C.S., Watenpaugh K.D., Wiley V.H.;
RT "Cation binding to the integrin CD11b I domain and activation model
RT assessment";
RL Structure 6:923-935(1998).
RN [12]
RP RX 3D-STRUCTURE MODELING OF 17-616.
RP RX MEDLINE=9826734; PubMed=9560195;
RA Oxvig C., Springer T.A.;
RT "Experimental support for a beta-propeller domain in integrin alpha-
RT subunits and a calcium binding site on its lower surface";
RL Proc. Natl. Acad. Sci. U.S.A. 95:4870-4875(1998).
CC -1- FUNCTION: INTEGRIN ALPHA-M/BETA-2 IS IMPLICATED IN VARIOUS
CC ADHESIVE INTERACTIONS OF MONOCYTES, MACROPHAGES AND GRANULOCYTES
CC AS WELL AS IN MEDIATING THE UPTAKE OF COMPLEMENT-COATED PARTICLES.
CC IT IS IDENTICAL WITH CR-3, THE RECEPTOR FOR THE IC3B FRAGMENT OF
CC THE THIRD COMPLEMENT COMPONENT. IT PROBABLY RECOGNIZES THE R-G-D
CC PEPTIDE IN C3B. INTEGRIN ALPHA-M/BETA-2 IS ALSO A RECEPTOR FOR
CC FIBRINOGEN, FACTOR X AND ICAM1. IT RECOGNIZES P1 AND P2 PEPTIDES
CC OF FIBRINOGEN GAMMA CHAIN.
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-M
CC ASSOCIATES WITH BETA-2.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND
CC GRANULOCYTES.
CC -1- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -1- SIMILARITY: Belongs to the integrin alpha chain family.
CC -1- SIMILARITY: Contains 1 VWFA domain.
CC -1- SIMILARITY: Contains 7 FG-GAP repeats.
CC -1- DATABASE: NAME=PROW; NOTE=CD guide CD11b entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd11b.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: J03925; AAA59544.1; -
CC EMBL: M18044; AAA59491.1; -
CC EMBL: J04145; AAA59903.1; -
CC EMBL: S52227; AAB24821.1; -
CC EMBL: S52152; AAB24821.1; JOINED.
CC EMBL: S52153; AAB24821.1; JOINED.
CC EMBL: S52154; AAB24821.1; JOINED.
CC EMBL: S52155; AAB24821.1; JOINED.
CC EMBL: S52157; AAB24821.1; JOINED.
CC EMBL: S52159; AAB24821.1; JOINED.
CC EMBL: S52161; AAB24821.1; JOINED.
CC EMBL: S52164; AAB24821.1; JOINED.
CC EMBL: S52165; AAB24821.1; JOINED.
CC EMBL: S52167; AAB24821.1; JOINED.
CC EMBL: S52169; AAB24821.1; JOINED.
CC EMBL: S52170; AAB24821.1; JOINED.
CC EMBL: S52173; AAB24821.1; JOINED.
CC EMBL: S52174; AAB24821.1; JOINED.
CC EMBL: S52180; AAB24821.1; JOINED.
CC EMBL: S52181; AAB24821.1; JOINED.
CC EMBL: S52184; AAB24821.1; JOINED.
CC EMBL: S52189; AAB24821.1; JOINED.
CC EMBL: S52191; AAB24821.1; JOINED.
CC EMBL: S52192; AAB24821.1; JOINED.
CC EMBL: S52203; AAB24821.1; JOINED.
CC EMBL: S52212; AAB24821.1; JOINED.
CC EMBL: S52213; AAB24821.1; JOINED.
CC EMBL: S52216; AAB24821.1; JOINED.
CC EMBL: S52219; AAB24821.1; JOINED.
CC EMBL: S52220; AAB24821.1; JOINED.
CC EMBL: S52221; AAB24821.1; JOINED.
CC EMBL: S52222; AAB24821.1; JOINED.
CC EMBL: S52226; AAB24821.1; JOINED.
CC EMBL: M76724; AAA58410.1; -
CC EMBL: M84477; AAA51960.1; -
CC PIR: A31108; RWHU1B.
CC PDB: 1A8X; 17-JUN-98.
CC PDB: 1BRO; 18-NOV-98.
CC PDB: 1BHQ; 18-NOV-98.
CC PDB: 1IDN; 25-NOV-98.
CC PDB: 1IDO; 01-AUG-96.
CC PDB: 1JUM; 11-JAN-97.
CC PDB: 1MIU; 07-AUG-02.
CC Genew; HGNC:6149; ITGAM.
CC MIM: 120980; -
CC GO: GO:0008305; C:integrin complex; TAS.
CC GO: GO:0004895; F:cell adhesion receptor activity; TAS.
CC GO: GO:0007155; P:cell adhesion; TAS.
CC InterPro: IPR000413; Integrin_alpha.
CC InterPro: IPR002035; VWF_A.
CC Pfam: PF01839; FG-GAP; 3.
CC Pfam: PF00357; Integrin_A; 1.
CC Pfam: PF00092; vwa; 1.
CC PRINTS; PRO1185; INTEGRINA.
CC PRINTS; PRO0453; VWFADOMAIN.
CC SMART; SM00191; int_alpha; 4.
CC SMART; SM00327; vwa; 1.
CC PROSITE; PS00242; INTEGRIN_ALPHA; 1.
CC PROSITE; PS50234; VWFA; 1.
CC Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
CC Signal; 3D-structure; Repeat; Magnesium; Calcium.
CC SIGNAL 1 16
CC CHAIN 17 1152 INTEGRIN ALPHA-M.
CC DOMAIN 17 1104 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 1105 1128 POTENTIAL.
CC DOMAIN 1129 1152 CYTOPLASMIC (POTENTIAL).
CC

Query Match		99.2%; Score 5836.5; DB 1; Length 1152;	
Best Local Similarity		99.1%; Pred. No. 0;	
Matches 1127; Conservative		7; Mismatches	2; Indels 1; Gaps 1;
QY	1	FNLDTENAMTFQENARGFQGSVVOVGSRVVVGAQEIIVAAQNRGLVOCYDSTGSCPEI 60	
DB	17	FNLDTENAMTFQENARGFQGSVVOVGSRVVVGAQEIIVAAQNRGLVOCYDSTGSCPEI 76	
QY	61	RLQVPEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVKGCLFGLFGLNLRQOQPK 120	
DB	77	RLQVPEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVKGCLFGLFGLNLRQOQPK 136	
QY	121	FPBALRGCCQEDSDIAFLIDGSGSIIPHDFRMEKWSVMEOLKSKTKLFLSMQYSEF 180	
DB	137	FPBALRGCCQEDSDIAFLIDGSGSIIPHDFRMEKWSVMEOLKSKTKLFLSMQYSEF 196	
QY	181	RIHFTFKFQNNPNRSLKIPITQLLGRTHRTATGLRKVVRELFNITNGARKNAKIFLFL 240	
DB	197	RIHFTFKFQNNPNRSLKIPITQLLGRTHRTATGLRKVVRELFNITNGARKNAKIFLVI 256	
QY	241	TGEXFGDPLGYEDVIPEDRGVIRYVIGVDGDAFRSEKSRQELNTVASKPPRDHVFOIN 300	
DB	257	TGEXFGDPLGYEDVIPEDRGVIRYVIGVDGDAFRSEKSRQELNTVASKPPRDHVFOIN 316	
QY	301	NFEALKTIONQLREKIFAETGQTGSSSFHEMSQEGFSAAITNGPLLSVTGVDWAG 360	
DB	317	NFEALKTIONQLREKIFAETGQTGSSSFHEMSQEGFSAAITNGPLLSVTGVDWAG 376	
QY	361	GVFLTSKEKSTFINNTRVDSMDNAYLGYYAAAILNRVQSLVLGAPRYQHIGLVAMPR 420	
DB	377	GVFLTSKEKSTFINNTRVDSMDNAYLGYYAAAILNRVQSLVLGAPRYQHIGLVAMPR 436	
QY	421	QNTGWMESNANVKGTQIGAYFGASLCSDVDNGSTDLVLGAPHYEQTRGQSVQVCL 480	
DB	437	QNTGWMESNANVKGTQIGAYFGASLCSDVDNGSTDLVLGAPHYEQTRGQSVQVCL 496	
QY	481	PRQGRARWQCDVLYGEOQPGRGFAALTVLGVDNGDKLTDVAICAPGEEDNRGAVLYF 540	
DB	497	PRG-RARWQCDVLYGEOQPGRGFAALTVLGVDNGDKLTDVAICAPGEEDNRGAVLYF 555	
QY	541	HGTSGSGISPSHSQRIAGSKLPRLOYFGOSLGGGDLTMDGLVDLTWGAQCHVLLRSQ 600	
DB	556	HGTSGSGISPSHSQRIAGSKLPRLOYFGOSLGGGDLTMDGLVDLTWGAQCHVLLRSQ 615	
QY	601	PVLRVKAIMFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660	
DB	616	PVLRVKAIMFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 675	
QY	661	YDLALDSGRPHSRVFNETKSTRQTQVIGLQTQTCETLKLQLPNCIEDPVSPIVRLNF 720	
DB	676	YDLALDSGRPHSRVFNETKSTRQTQVIGLQTQTCETLKLQLPNCIEDPVSPIVRLNF 735	
QY	721	SLVGTPLSAFGNLRPVLAEDAQLPTALPPFEKNCNDNI CODDLSITFSFMSLDCLVVG 780	
DB	736	SLVGTPLSAFGNLRPVLAEDAQLPTALPPFEKNCNDNI CODDLSITFSFMSLDCLVVG 795	
QY	781	GPREFNVTVVNDGEDSVRTQVTFPPFLDLKYRVSTLQNRQSRWRLACESASSTEV 840	
DB	796	GPREFNVTVVNDGEDSVRTQVTFPPFLDLKYRVSTLQNRQSRWRLACESASSTEV 855	
QY	841	SGALKSTCSINHPFPENSEVTENITPDVDSKASLGKLLKANVTSENMPRTNKTEP 900	
DB	856	SGALKSTCSINHPFPENSEVTENITPDVDSKASLGKLLKANVTSENMPRTNKTEP 915	
QY	901	QLELPVKAVVMVTVSHGVSSTKYLNTASNTSRVWQHQQYQVSNLQGSPLSLFLVLPV 960	
DB	916	QLELPVKAVVMVTVSHGVSSTKYLNTASNTSRVWQHQQYQVSNLQGSPLSLFLVLPV 975	
QY	961	RLNQTIVMDRPQVTVSENLSSTCHTKERLPSHSDFLAELKAPVNCIAVCORIQCDIP 1020	
DB	976	RLNQTIVMDRPQVTVSENLSSTCHTKERLPSHSDFLAELKAPVNCIAVCORIQCDIP 1035	
QY	1021	PFQIQEFNATLKNLSFDWIKTSHNHLIVSTABILFNDSVFTLLPGQAFVRSQTET 1080	
DB	1036	PFQIQEFNATLKNLSFDWIKTSHNHLIVSTABILFNDSVFTLLPGQAFVRSQTET 1095	
QY	1081	KVEPFEVNPPLVIVGSSVGLLLALITAAALYKLGFFKQYKDMWSSGGPFGAEPQ 1137	
DB	1096	KVEPFEVNPPLVIVGSSVGLLLALITAAALYKLGFFKQYKDMWSSGGPFGAEPQ 1152	
RESULT 2			
ITAM MOUSE			
ID	ITAM	MOUSE	STANDARD; PRT; 1153 AA.
AC	P0555;		
DT	01-NOV-1988	(Rel. 09, Created)	
DT	01-FEB-1991	(Rel. 17, Last sequence update)	
DT	28-FEB-2003	(Rel. 41, Last annotation update)	
DE	Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor MO1).		
GN	ITGAM.		
OS	Mus musculus (Mouse);		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	SEQUENCE FROM N.A.		
RP	MEDLINE=88312584; PubMed=3044779;		
RX	Pyela R.;		
RT	"Amino acid sequence of the murine Mac-1 alpha chain reveals homology with the integrin family and an additional domain related to von Willebrand factor.";		
RL	EMBO J. 7:1371-1378(1988).		
RN	SEQUENCE OF 11-45 FROM N.A.		
RP	STRAIN=BALB/c; TISSUE=Spleen;		
RC	MEDLINE=85287312; PubMed=2942940;		
EX	Sastre L., Roman J.M., Teplow D.B., Dreyer W.J., Gee C.E.,		
RA	Larson R.S., Roberts T.W., Springer T.A.;		
RT	"A partial genomic DNA clone for the alpha subunit of the mouse complement receptor type 3 and cellular adhesion molecule Mac-1.";		
RT	Proc. Natl. Acad. Sci. U.S.A. 83:5644-5648(1986).		
RL	Nature 314:540-542(1985).		
RN	FUNCTION: INTEGRIN ALPHA-M/BETA-2 IS IMPLICATED IN VARIOUS ADHESIVE INTERACTIONS OF MONOCYTES, MACROPHAGES AND GRANULOCYTES AS WELL AS IN MEDIATING THE UPTAKE OF COMPLEMENT-COATED PARTICLES. IT IS IDENTICAL WITH CR-3, THE RECEPTOR FOR THE IC3B FRAGMENT OF THE THIRD COMPLEMENT COMPONENT. IT PROBABLY RECOGNIZES THE R-G-D PEPTIDE IN C3B. INTEGRIN ALPHA-M/BETA-2 IS ALSO A RECEPTOR FOR FIBRINOGEN, FACTOR X AND ICAM1. IT RECOGNIZES P1 AND P2 PEPTIDES OF FIBRINOGEN GAMMA CHAIN. ALPHA-M/BETA-2 PLAY A CRITICAL ROLE IN MAST CELL DEVELOPMENT AND IN IMMUNE COMPLEX-MEDIATED GLOMERULONEPHRITIS. MICE EXPRESSING A NULL MUTATION OF THE ALPHA-M SUBUNIT GENE DEMONSTRATE INCREASE IN NEUTROPHIL ACCUMULATION, IN RESPONSE TO AN IMPAIRED DEGRADATION AND PHAGOCYTOSIS, EVENTS THAT APPARENTLY ACCELERATE APOPTOSIS IN NEUTROPHILS. THESE MICE DEVELOP OBESITY.		
CC	!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN. ALPHA-M ASSOCIATES WITH BETA-2.		
CC	!- SUBCELLULAR LOCATION: Type I membrane protein.		
CC	!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND GRANULOCYTES.		
CC	!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VFMA DOMAIN. INTEGRINS WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.		
CC	!- SIMILARITY: Belongs to the integrin alpha chain family.		
CC	!- SIMILARITY: Contains 1 VFMA domain.		
CC	!- SIMILARITY: Contains 7 FG-GAP repeats.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

CC EMBL; X07640; CAA30479.1; -;
 CC EMBL; M14293; AAA39484.1; -;
 CC PIR; S00551; S00551.
 CC HSP; P11215; IABX.
 CC MGSP; MGI196607; Itgam.
 CC InterPro; IPR000413; Integrin_alpha.
 CC InterPro; IPR002035; VWF_A.
 CC Pfam; PF01839; FG-GAP; 3.
 CC Pfam; PF00357; Integrin_A; 1.
 CC Pfam; PF00092; vwa; 1.
 CC PRINTS; PR01185; INTEGRINA.
 CC PRINTS; PR00453; VWFADOMAIN.
 CC SMART; SM00191; Int_alpha; 5.
 CC SMART; SM00327; VWA; 1.
 CC PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 CC PROSITE; PS00234; VWA; 1.
 CC Integrin; Call adhesion; Receptor; Glycoprotein; Transmembrane;
 KW Signal; Calcium; Repeat.
 FT SIGNAL 1
 FT CHAIN 17 1153
 FT DOMAIN 17 1105
 FT TRANSMEM 1106 1129
 FT DOMAIN 1130 1153
 FT REPEAT 31 84
 FT REPEAT ? ?
 FT DOMAIN 164 350
 FT REPEAT 337 400
 FT REPEAT 401 452
 FT REPEAT 454 515
 FT REPEAT 517 575
 FT REPEAT 580 632
 FT CA_BIND 465 473
 FT CA_BIND 529 537
 FT CA_BIND 592 600
 FT SITE 1132 1136
 FT DISULFID 66 73
 FT DISULFID 105 123
 FT DISULFID 654 711
 FT DISULFID 770 776
 FT DISULFID 999 1023
 FT DISULFID 1028 1033
 FT CARBOHYD 58 58
 FT CARBOHYD 86 86
 FT CARBOHYD 391 391
 FT CARBOHYD 696 696
 FT CARBOHYD 734 734
 FT CARBOHYD 772 772
 FT CARBOHYD 801 801
 FT CARBOHYD 881 881
 FT CARBOHYD 907 907
 FT CARBOHYD 941 941
 FT CARBOHYD 980 980
 FT CARBOHYD 984 994
 FT CARBOHYD 1022 1022
 FT CARBOHYD 1045 1045
 FT CARBOHYD 1051 1051
 FT CARBOHYD 1076 1076
 FT CARBOHYD 1153 1153
 SQ SEQUENCE 1153 AA; 127480 MW; 178DB988AECB0343 CRC64;

Query Match 75.88; Score 4460; DB 1; Length 1153;
 Best Local Similarity 73.88; Pred. No. 1.8e-290;
 Matches 840; Conservative 146; Mismatches 150; Indels 2; Gaps 2;
 1 FNLDTEHMTFQENARGFGQSVVQLQGSRRVVGAPQEIIVAAVNRGSLYQCDYSTGSCPEI 60
 17 FNLDTEHMTFQENARGFGQSVVQLQGSRRVVGAPQEIIVAAVNRGSLYQCDYSTGSCPEI 76

QY 61 RLOVPEAVNMKSLGLSLAATTSPOLLACGPTVHOTCSNTYVKGCLPGLSGNLQQPOK 120
 DB 77 PIQVPEAVNMKSLGLSLAATTSPOLLACGPTVHOTCSNTYVKGCLPGLSGNLQQPOK 136
 QY 121 PFEALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKEWVSTVMBQKSKSTLPGLMYSSEF 180
 DB 137 PFEALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKEWVSTVMBQKSKSTLPGLMYSSEF 196
 QY 181 RIHFTFK3QNNPNPRSLIKPITOLLGRTHATGKRVKVRRELFNITNGARKNAFKILPLL 240
 DB 197 RIHFTFNDPKRNPSPRSHVSPKIQNGRTKTASGIRKVVVRELFNITNGARKNAFKILVVI 256
 QY 241 TDGKFGDPLGYEDVIPELDRGVIRYVIGVDAPFRSKSQELNTAVSKPPRDRHVFQIN 300
 DB 257 TDGKFGDPLGYEDVIPELDRGVIRYVIGVDAPFRSKSQELNTAVSKPPRDRHVFQIN 316
 QY 301 NFEALKTIONQLREKIPALEGTQTCSSSFHEHMSQEGFSAITNSGPIISVGVSDYDAG 360
 DB 317 NFEALKTIONQLREKIPALEGTQTCSSSFHEHMSQEGFSAITNSGPIISVGVSDYDAG 376
 QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGZAAAIILNRVQSLVGLGAPRYCHIGLVAMFR 420
 DB 377 GAFLYTSKDKVTFINTTRVDSMDNDAYLGZAAAIILNRVQSLVGLGAPRYCHIGLVAMFR 436
 QY 421 QNTGAWESNANVKGTQIGAYFGASLCSDVDNSGSTDILVIGAPHYYOTRGQVSVCPD 480
 DB 437 ENFGTWBPHSTIKSGQISYFGASLCSDVDNSGSTDILVIGAPHYYOTRGQVSVCPD 496
 QY 481 PRGQARWOCDAVLYGBOGQPMRGFAALTVDLVGNGDKLTDVAIGAPGEEENRGAVILF 540
 DB 497 PRG-PARWQCEALLHGDQHPWGRFGAALTVDLVGNGDKLTDVAIGAPGEEENRGAVILF 555
 QY 541 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLGSGQDLTMDGLVDLTVGAQHVLRLSQ 600
 DB 556 YGASTASLSASHSHRIIGAHFSPGLQYFGQSLGSGQDLTMDGLVDLTVGAQHVLRLSQ 615
 QY 601 PVLRYKALMEFNPREVARNVPCNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT 660
 DB 616 PVLRYKALMEFNPREVARNVPCNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT 675
 QY 661 YDLALDGRPHSRVAFNETKSTRTQTVLGTQTCETLKLQLPNCIEDPVPVILRLNF 720
 DB 676 YDLALDGRPHSRVAFNETKSTRTQTVLGTQTCETLKLQLPNCIEDPVPVILRLNF 735
 QY 721 SLVGTPLSAFGLNLRVLAEDAQRLELTALPPEKNGNDNICODDLSITFSEMSLCLVVG 780
 DB 736 TLVGTPLSAFGLNLRVLAEDAQRLELTALPPEKNGNDNICODDLSITFSEMSLCLVVG 795
 QY 781 GPREFNVTVVRNDGEDSYRTQTVLGTQTCETLKLQLPNCIEDPVPVILRLNF 839
 DB 796 GPREFNVTVVRNDGEDSYRTQTVLGTQTCETLKLQLPNCIEDPVPVILRLNF 855
 QY 840 VSGALKSTSCSINHPFPENSEVTNITFDVDSKASLGNKLLKANTYSENMPRTNKT 899
 DB 856 GHGALKSTSCSINHPFPENSEVTNITFDVDSKASLGNKLLKANTYSENMPRTNKT 915
 QY 900 FQLELPVXYAVVWVTSVSHGVSTKYLNFTASENTSRVMOHQVQVSNLQORSLPISLVFLVP 959
 DB 916 FQLELPVXYAVVWVTSVSHGVSTKYLNFTASENTSRVMOHQVQVSNLQORSLPISLVFLVP 975
 QY 960 VRLNQTIVWDRPQVTPSENLSSTCHTKERLPSHSDFLAELRKA PVVNCISIAVCRIQCDI 1019
 DB 976 VQINNVTWDRPQVTPSENLSSTCHTKERLPSHSDFLAELRKA PVVNCISIAVCRIQCDI 1035
 QY 1020 PPFQIQEBEPNATLKNLSFDDWYIKTSHNELITVSTAEILFNDSTVTLPGQAFVRSQTE 1079
 DB 1036 PPFQIQEBEPNATLKNLSFDDWYIKTSHNELITVSTAEILFNDSTVTLPGQAFVRSQTE 1095
 QY 1080 TKVEPPEVNPPLPLTVGSSVGLLLALITLALYKLGFFKQYKDMGEGGPPGPEQ 1137
 DB 1096 TKVEPPEVNPPLPLTVGSSVGLLLALITLALYKLGFFKQYKDMGEGGPPGPEQ 1153

```
RESULT 3
IPAX_HUMAN
ID ITAX_HUMAN STANDARD; PRT; 1163 AA.
AC P20702;
DT 01-FEB-1991 (Rel. 17, Created)
DE 01-FEB-1991 (Rel. 17, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Integrin alpha-X precursor (Leukocyte adhesion glycoprotein p150,95
DE alpha chain) (Leukocyte adhesion receptor p150,95) (CD11c) (Leu M5).
GN ITGAX OR CD11c.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88166645; PubMed=3327687;
RA Corbi A.L., Miller L.J., O'Connor K., Larson R.S., Springer T.A.;
RT "cDNA cloning and complete primary structure of the alpha subunit of
RT a leukocyte adhesion glycoprotein, p150,95."
RL EMBO J. 6:4023-4028(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90153906; PubMed=2303426;
RA Corbi A.L., Garcia-Aguilar J., Springer T.A.;
RT "Genomic structure of an integrin alpha subunit, the leukocyte
RT p150,95 molecule."
RL J. Biol. Chem. 265:2782-2788(1990).
RN [3]
RP ERRATUM.
RA Corbi A.L., Garcia-Aguilar J., Springer T.A.;
RL J. Biol. Chem. 265:12750-12751(1990).
RN [4]
RP SEQUENCE OF 20-43.
RX MEDLINE=87167596; PubMed=3549901;
RA Miller L.J., Wiebe M., Springer T.A.;
RT "Purification and alpha subunit N-terminal sequences of human Mac-1
RT and p150,95 leukocyte adhesion proteins."
RL J. Immunol. 138:2381-2383(1987).
CC -!- FUNCTION: INTEGRIN ALPHA-X/BETA-2 IS A RECEPTOR FOR FIBRINOGEN. IT
CC RECOGNIZES THE SEQUENCE G-P-R IN FIBRINOGEN. IT MEDIATES CELL-CELL
CC INTERACTION DURING INFLAMMATORY RESPONSES. IT IS ESPECIALLY
CC IMPORTANT IN MONOCYTE ADHESION AND CHEMOTAXIS.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-X
CC ASSOCIATES WITH BETA-2.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND
CC GRANULOCYTES.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD11c entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd11c.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M81695; AA59180.1; -;
DR EMBL; Y00093; CAAG6283.1; -;
DR EMBL; M29165; -; NOT ANNOTATED CDS.
DR EMBL; M29487; AA51620.1; ALT SEQ.
DR EMBL; M29482; AA51620.1; JOINED.
DR EMBL; M29483; AA51620.1; JOINED.
DR EMBL; M29484; AA51620.1; JOINED.
DR EMBL; M29485; AA51620.1; JOINED.
```

```
DR EMBL; M29486; AA51620.1; JOINED.
DR EMBL; A36584; RWHULC.
DR PDB; 1N3Y; 18-FEB-03.
DR Genew; HGNC:6152; ITGAX.
DR MIM; 151510; -.
DR GO; GO:0008305; C: integrin complex; TAS.
DR GO; GO:0004895; F: cell adhesion receptor activity; TAS.
DR GO; GO:0004872; F: receptor activity; TAS.
DR GO; GO:0007155; P: cell adhesion; TAS.
DR GO; GO:0007397; P: histogenesis and organogenesis; TAS.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWFA.
DR Pfam; PF01839; FG-GAP; 1.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00131; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWFA; 1.
DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Signal; Magnesium; Calcium; Repeat; 3D-structure.
FT SIGNAL 1 19
FT CHAIN 20 1163 INTEGRIN ALPHA-X.
FT DOMAIN 20 1107 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1108 1128 POTENTIAL.
FT DOMAIN 1129 1163 CYTOPLASMIC (POTENTIAL).
FT REPEAT 34 87 FG-GAP 1.
FT REPEAT ? ? FG-GAP 2.
FT DOMAIN 165 351 VWFA.
FT REPEAT ? ? FG-GAP 3.
FT REPEAT 402 453 FG-GAP 4.
FT REPEAT 455 517 FG-GAP 5.
FT REPEAT 518 576 FG-GAP 6.
FT REPEAT 581 633 FG-GAP 7.
FT CA_BIND 466 474 POTENTIAL.
FT CA_BIND 530 538 POTENTIAL.
FT SITE 593 601 POTENTIAL.
FT SITE 1131 1135 GFFKR MOTIF.
FT DISULFID 69 76 BY SIMILARITY.
FT DISULFID 108 126 BY SIMILARITY.
FT DISULFID 655 712 BY SIMILARITY.
FT DISULFID 771 777 BY SIMILARITY.
FT DISULFID 848 863 BY SIMILARITY.
FT DISULFID 998 1022 BY SIMILARITY.
FT DISULFID 1027 1032 BY SIMILARITY.
FT CARBOHYD 61 61 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 89 89 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 697 697 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 735 735 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 899 899 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 939 939 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1050 1050 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 490 490 G -> A (IN REF. 2).
FT CONFLICT 756 756 L -> D (IN REF. 2).
SQ SEQUENCE 1163 AA; 127885 MW; 6C4B19CC3F62AA73 CRC64;
```

Query Match 58.8%; Score 3459; DB 1; Length 1163;

Best Local Similarity 60.9%; Pred. No. 2e-223;

Matches 687; Conservative 142; Mismatches 294; Indels 6; Gaps 4;

QY 1 FNLDTEAMTQENARGFGQSVVQLQGSRRVVVGAPQEIIVAAVNRGSLYQCDYSTGSCPEI 60

DB 20 FNLDTEELTAFRVDSAGFSDSVVQVANSVWVGAPQKITAANQTGLYQCGYSTGACEFI 79

QY 61 RIQVPEAVNMSLGLSLAATSPQLLACGTVVQTCSEYVYVXGLCLFGLSNLRQOQK 120

DB 80 GLQVPEAVNMSLGLSLASTTSPQLACGTVVHHECGRNMVLTGLCFLGPT--QLTOR 137

QY 121 FPEALRGCPQSDSDIAFLIDGSGSIIPHDFRMRKEWSTVMEQAKKSKTLFSLMOYSBEF 180

```

Db 138 LPVSROECPQEQDIYFLIDGSGSISRNFATMNFVRAVISQFQSPSTQFSLMQFSNKE 197
Qy 181 RIHFTKEFQNNPNRSLIKPIQLLGRTHATGLRKRVRELPNIYNGARXNAFKILFL 240
Db 198 QTHFTFEFEERTNPISLLASVHLQLOQFTYTAIONVVRHLFHASYGARRDAIKILIVI 257
Qy 241 TDGKFGDPLGYEDVLPEDREGVIRVIGVGAFAFSEKSRQELNTVASKPRDHVFCQN 300
Db 258 TDGKKGDSLDYKDVLPADAAAGIYAIGVGLAFQNEUSWKLNDIAKPSQEHIFKVE 317
Qy 301 NFALKTIQNLREKIFAIEGTOTGSSSPHEHMSQEGFSAALITNGPILLSVGVSDWAG 360
Db 318 DFDALXDIQNLREKIFAIEGTETITSSSPSELEMAQEGFSAVFTDGPVLGAGSFTWSG 377
Qy 361 GVFLYTSKESKSTINTRVDSNDAYLGVAAAILRNVRQSLVLCAPRYOHLGLVAMER 420
Db 378 GAFLYPPNMSPTINNSQENVNDRDSYLGSTELALWGVQSLVLCAPRYOHTGKAVIT 437
Qy 421 QNTGMESNANVGTQIGAYFGASLCSVDVDSNGSDTLVLIGAPHYYEOTRGQSVVCEL 480
Db 438 QVSRQWRMKAENVGTQIGSYFGASLCSVDVDTGSDTLVLIGAPHYYEOTRGQSVVCEL 497
Qy 481 PRGORARQCDVLYGEGQGPWCFGAALTVLGDVNGDKLTQVAIGAPGEDNRRGAVILF 540
Db 498 PRGWR-RWQCDVLYGEGQGPWCFGAALTVLGDVNGDKLTQVAIGAPGEDNRRGAVILF 556
Qy 541 HGTSGGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 557 HGVLPSPISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTQDGLVDLAVGARGQVLLLR 616
Qy 601 PVLRYKALMFNPREVARNVPCNDVNVKKEAGEVRVCLHVOKSTEDRLRQEQISVVT 660
Db 617 PVLWGVGVMQFIPABIPRPAFECREQVVSQTLVQSNICLYIDKRSKLLGSHDLQSSVT 676
Qy 661 YDLALDGRPHRAVENETKNSRRTQVTLGLTQTCETLKLQPCNIEPVPVSVILRLNF 720
Db 677 LDALDGRPLSPRATQETKNSLSRVRLGLAKHCENFNLLPSCVEDSVTPIITLRNF 736
Qy 721 SIVGTPLSARCNLPVLAEDAORLFRALPPEKNCNDNIQCDDLSITFSFMSLCLVVG 780
Db 737 TLVCKPLAFNLRLMLAALAQRYFTASLPFEKNCGADHICQDNLGISFSPGLKSLVG 796
Qy 781 GPREFNVTVVRNDGDSYRTVTFPPPLDLSYKRVSTLQNSQSRSLWKLACSSASTEV 840
Db 797 SNLELNAEVWVNDGDSYCTITFSPHAGLSRYVAEGKQQLASLHLTCDSPVGV-- 854
Qy 841 SGALKSTSCSINPIFENSEVTNFTFDVDSKASLGNKLLKANTSENNNMPRTKTEF 900
Db 855 SQGTWSTSCRINHLIFRGGAQITFLATFDVSPKAVLGDRLLLTANYSENNTPTKSTTF 914
Qy 901 QLELPVKYAVVMVTSBGVSTKYLNFPTAS-ENTSRVMOHOYQVSNLQORSLPISLFLVP 959
Db 915 QLELPVKYAVTVVSSEHQTKYLNFSESEKESHVAMHRYQVNNLQORDLPVSNFWVP 974
Qy 960 VRLNQTVIMDRPQVTFSENLSSTCHTKERLPSHDSFLAELRKAPVWNCSTAVCQICDI 1019
Db 975 VELNQAVMMDVEVSHPNPQSLRCSSEKIAPPASDFLAHTQKNPVLDCSIAGCLRFCDV 1034
Qy 1020 PPGIOEFNATLKNLSFMYIKTSHNLLIYSTAILFENDSVFTLLPQGAFFVSQTE 1079
Db 1035 PFSVQBELDFTLKNLSFGWVQIOLQKYSVSVSAEITDTTSVQLPQGEAPMRAQT 1094
Qy 1080 TKVEPFEVNPPLPIVGVSSVGGILLALITAAALYKLGFFKRYQKQKMMSE 1128
Db 1095 TVLEKRYVHNPTPLIVGVSSIGILLALITAVLYKVGFFKRYQKEMWEE 1143

RESULT 4
ITAD HUMAN
ID ITAD HUMAN STANDARD; PRT: 1162 AA.
AC Q13349; Q15575; Q15576;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)

```

```

DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-D precursor (Leukointegrin alpha D) (CD11d) (ADB2).
GN ITGAD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen.
RX MEDLINE=9611956; PubMed=8777714;
RA Van der Vliet M., Le Trong H., Wood C.L., Moore P.F., St John T.,
RA Staunton D.E., Gallatin W.M.;
RT "A novel leukointegrin, alpha d beta 2, binds preferentially to ICAM-
RT 3."
RL Immunity 3:683-690 (1995).
RN [2]
RP SEQUENCE OF 1-235 FROM N.A.
RX MEDLINE=20187620; PubMed=10722744;
RA Noti J.D., Johnson A.K., Dillon J.D.;
RT "Structural and functional characterization of the leukocyte integrin
RT gene CD11d. Essential role of Sp1 and Sp3."
RL J. Biol. Chem. 275:8959-8969 (2000).
RN [3]
RP SEQUENCE OF 457-537; 571-602; 633-666; 788-834 AND 910-1125 FROM N.A.
RX MEDLINE=96257236; PubMed=8666289;
RA Wong D.A., Davis B.M., LeBeau M., Springer T.A.;
RT "Cloning and chromosomal localization of a novel gene encoding a human
RT beta 2-integrin alpha subunit."
RL Gene 171:291-294 (1996).
RN [4]
RP INTERACTION WITH VCAM1.
RX MEDLINE=99059842; PubMed=9841932;
RA Grayson M.H., Van der Vliet M., Sterbinsky S.A., Michael Gallatin W.,
RA Hoffman P.A., Staunton D.E., Bochner B.S.;
RT "alpha2beta2 integrin is expressed on human eosinophils and functions
RT as an alternative ligand for vascular cell adhesion molecule 1
RT (VCAM-1)."
RL J. Exp. Med. 188:2187-2191 (1998).
RN [5]
RP INTERACTION WITH VCAM1.
RX MEDLINE=99370002; PubMed=10438935;
RA Van der Vliet M., Crowe D.T., Hoekstra D., Vazeux R., Hoffman P.A.,
RA Grayson M.H., Bochner B.S., Gallatin W.M., Staunton D.E.;
RT "The leukocyte integrin alpha D beta 2 binds VCAM-1: evidence for a
RT binding interface between I domain and VCAM-1."
RL J. Immunol. 163:1984-1990 (1999).
CC -!- FUNCTION: INTEGRIN ALPHA-D/BETA-2 IS A RECEPTOR FOR ICAM3 AND
CC CLEARING LIPOPROTEINS FROM PLAQUES AND IN PHAGOCYTOSIS OF BLOOD-
CC BORNE PATHOGENS, PARTICULATE MATTER, AND SENESECENT ERYTHROCYTES
CC FROM THE BLOOD.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-D
CC ASSOCIATES WITH BETA-2.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: EXPRESSED MODERATELY ON MYELOMONOCYTIC CELL
CC LINES AND SUBSETS OF PERIPHERAL BLOOD LEUKOCYTES AND STRONGLY ON
CC TISSUE-SPECIALIZED CELLS, INCLUDING MACROPHAGES FOAM CELLS WITHIN
CC ATHEROSCLEROTIC PLAQUES, AND ON SPLENIC RED PULP MACROPHAGES.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: belongs to the integrin alpha chain family.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

DR EMBL; U37028; AAB38547.1; -
 DR EMBL; U40274; AAB60634.1; -
 DR EMBL; U40275; AAB60635.1; -
 DR EMBL; U40276; AAB60636.1; -
 DR EMBL; U40277; AAB60637.1; -
 DR EMBL; U40278; AAB60638.1; -
 DR EMBL; U40279; AAB60639.1; -
 DR EMBL; U40278; AAB60638.1; JOINED.
 DR EMBL; AF187881; AAF62875.1; -
 DR HSSP; F11215; IABX.
 DR HSSP; F11215; IABX.
 DR MIM; 602453; -
 DR GO; GO:0008305; C:integrin complex; TAS.
 DR GO; GO:0004895; P:cell adhesion receptor activity; TAS.
 DR GO; GO:0015337; P:cell-cell adhesion; NAS.
 DR GO; GO:0007160; P:cell-matrix adhesion; NAS.
 DR GO; GO:0006955; P:immune response; NAS.
 DR InterPro; IPR000413; Integrin_alpha.
 DR InterPro; IPR002035; VWF A.
 DR Pfam; PF01839; FG-GAP; 3.
 DR Pfam; PF00357; Integrin_A; 1.
 DR Pfam; PF00092; vwa; 1.
 DR PRINTS; PRO1185; INTEGRIN_A.
 DR PRINTS; PRO0453; VWFADOMAIN.
 DR SMART; SM00191; Int_alpha; 4.
 DR SMART; SM00327; VWA; 1.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 DR PROSITE; PS00234; VWA; 1.
 KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 KW Signal; Repeat; Calcium;
 KW Magnesium.
 FT SIGNAL 1 17
 FT CHAIN 18 1162
 FT DOMAIN 18 1100
 FT TRANSMEM 1101 1124
 FT DOMAIN 1125 1162
 FT REPEAT 32 85
 FT REPEAT ? ?
 FT DOMAIN 150 332
 FT REPEAT 350 400
 FT REPEAT 401 452
 FT REPEAT 454 516
 FT REPEAT 518 576
 FT REPEAT 581 633
 FT CA_BIND 465 473
 FT CA_BIND 530 538
 FT CA_BIND 593 601
 FT SITE 1127 1131
 FT DISULFID 67 74
 FT DISULFID 106 124
 FT DISULFID 655 710
 FT DISULFID 769 775
 FT DISULFID 846 861
 FT DISULFID 994 1018
 FT DISULFID 1023 1028
 FT CARBOHYD 59 1028
 FT CARBOHYD 87 97
 FT CARBOHYD 99 99
 FT CARBOHYD 391 391
 FT CARBOHYD 691 691
 FT CARBOHYD 733 733
 FT CARBOHYD 873 873
 FT CARBOHYD 957 957
 FT CARBOHYD 1046 1046
 FT CONFLICT 500 500
 FT CONFLICT 515 518
 FT CONFLICT 825 825
 FT CONFLICT 984 984
 FT SEQUENCE 1162 AA; 126885 MW; P296A1A35455D77D CRC64;

Query Watch 57.8%; Score 3401; DB 1; Length 1162;

Best Local Similarity 59.3%; Pred. No. 1.5e-219;

Matches 669; Conservative 167; Mismatches 285; Indels 8; Gaps 6;

QY 1 FNLDTEAMTFQENARCGSGVVOLOCSRVVVGAPQBIIVAAORGSLYQCDYSTGSCPEI 60
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 77
 18 FNLDVEEPTTFQEDAGGFGOSVWQFGSRLVVGAPLEVAANQGRLYDCAAAAGMCOPI 77
 QY 61 RLOVPVEAVNMSLGLSLAAATPPQLLACGPTVHQTCSNTYVXGLCFGLFSGNLRQOPK 120
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 136
 78 PLHIRPEAVNMSLGLTLAASINGSRLLACGPTLHRCVGENSYSGSCULLGSRW-ELIQT 136
 QY 121 FPEALRCQPOEDSOIAFLIDSGSIIIPHDPRMKEWSTVWEOLKSKTLFSLMOYSEEP 180
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 196
 137 VPDATPECPHQEMDIVFLIDGSGSIDQNDQNMKGFGVQAVMGQFEGTDTLTFALQVSWLL 196
 QY 181 RIHTFKEFQNNPNRSLIKPIITOLLGRTHATGLRKVVRELFINITNGARKNAKIFLL 240
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 256
 197 KIHTFTQFRTSPSQOSLVDPVQLKGLTPTATGILVTVTOLPHKNGARKSAKILIVI 256
 QY 241 TDGKFGDPLGYBDVIPDLREGVIRVIVGVGDAPRSEKSRQBLMTVASKPRPRHVFQIN 300
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 316
 257 TDGQKYDPLEYSDVIPQAEKAGIRYAIQVGHAFQGPARTARQELMTISSAPPQDHVKVD 316
 QY 301 NFEALKTIONOLREKIPIAIBGTOTGSSSSPHEMSQCFSAATISNGPLISTVGSYDWAG 360
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 376
 317 NFAALGSIQKQKELIYAVEGTQSRASSQFHEMSQEGFSTALIMDGLFLGAVGSPSWG 376
 QY 361 GVFLYTSKEKSTFINMTREVDMDNDAYLGAAAIILNRVQSLVGLGAPRYOHIGLVAMFR 420
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 436
 377 GAFLYPPNMSPTFINMSQENVMDRDSYLVGYSTELALWQVQLVGLGAPRYOHTGKAVIF 436
 QY 421 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVILGAPHYTQETGGGVSVCP 480
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 496
 437 QVSCWRKKAETGTQISYFGASLCSVDVDSNGSTDLVILGAPHYTQETGGGVSVCP 496
 QY 481 PRGORARWQCDVILYGEQGPWGEFGAALTALVILGVNVDGKLTDDVAIGAPGEEDNRGAVYLP 540
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 556
 497 PRGORVQWQCDVILYGEQGPWGEFGAALTALVILGVNVDGKLTDDVAIGAPGEEDNRGAVYLP 556
 QY 541 HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLGGQDLTMDGLVDLTGAGHVLLESQ 600
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 616
 557 HGASESGISPSHSQRIAGSKLSPLQYFGQSLGGQDLTMDGLVDLTGAGHVLLESQ 616
 QY 601 PVLVKALMBENPREVARNVPEQNDQVVKGEAGEVERVCLHVQKSTRDRLESGOITSVVT 660
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 674
 617 PVLKVGVMRSPVEVAKAVYRCWEBEKSALAGDAITVCLTIQKSSLDQL--GDIQSSVR 674
 QY 661 YDLALDGRPHSRVAVFNETKNSRQTVGLTQCTBLTKQLPNCIEDPVSPIVLRINF 720
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 734
 675 PDLALDPGLTSLRAIFNETKNPTLTRBKTLGLGHCHETLKLLEDVVDVSPILHLNF 734
 QY 721 SLVGTPLSAFGLNRPVLAEDAQRLEPTALFPEKKNCGNDNICODDLSTIFPSFMSLDCLVVG 780
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 794
 735 SLVREPSPQNLPRVLAAGSQDLPTASLPFERKNCQDGLCEGDLVTLSPSGQLTVG 794
 QY 781 GPREFNVTVYRNDGDSYRTQVTFPPDLISYRKVSTLQNRQSRWRLACRSASSTEV 840
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 853
 795 SSLEINLVITVWVNAAGEDSYGTWVSLVYPAGLSHRRVSGAQKQPHQSALELACETV-PTED 853
 QY 841 SCALAKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 912
 854 SG-LRSSRCSVNHPIFHEGSGNGTIVTFDVSFKATLGRMLMRASSENKAKSSKATP 912
 QY 901 QLELPVKYAVVMVVTSHGVSTKYNF-TASENTRVMQHQYQVSNLQORSUPISLVFLVP 959
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 972
 913 QLELPVKYAVVMVVTSHGVSTKYNF-TASENTRVMQHQYQVSNLQORSUPISLVFLVP 972
 QY 960 VRLNQTVINDRPPQVTFSENLSSTCHTKERLPSSHSDFLAELAKAPVWNCIAVCQRIQCDI 1019
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 1030
 973 VLLNGVAVMDVWMBAPSQSL--PCVSEKPKPQHSDFLTQTSRSPMLDGSIDCLQFRCDV 1030
 QY 1020 PFFGIGREFNATLKGNSLFWYIKTSHNHLIIVSTABEILFNDSTVFTLLPGQAGAFVRSQTE 1079
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 1090
 1031 PPSVQEBELDTLKGNSLFGWVRETLOKVLVSVSAEITPDTSVYSQLPQGAFAFWAQWE 1090
 QY 1080 TKVBPFEVNPFLPLIVGSSVGGLLLLALLITLAALYKLGFEFRQYKDMMS 1128


```

FT CARBOHYD 65 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 89 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 188 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 649 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 670 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 726 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 730 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 862 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 885 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 897 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 1060 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 1071 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT VARSPLIC 954 Q -> QGVHGLVEMQTSKQILCRAPGAEHTVGAQGEPLC
FT PMSVSEAFNRINRAGPCR (in isoform 2).
FT /FTIG-VSP 002738.
FT R -> W (IN REF. 1 AND 2).
FT Y -> I (IN REF. 2).

CONFLICT 214
FT STRAND 660
FT STRAND 155
FT STRAND 164
FT TURN 166
FT HELIX 185
FT TURN 186
FT STRAND 191
FT STRAND 202
FT HELIX 208
FT HELIX 217
FT TURN 222
FT STRAND 229
FT HELIX 233
FT TURN 244
FT HELIX 247
FT TURN 250
FT TURN 253
FT STRAND 256
FT HELIX 274
FT TURN 277
FT STRAND 280
FT HELIX 288
FT HELIX 293
FT TURN 298
FT HELIX 300
FT HELIX 307
FT STRAND 311
FT TURN 318
FT HELIX 319
FT TURN 329
SQ SEQUENCE 1170 AA; 128819 MW; 39A7AF92EF286FC0 CRC64;

Query Match
Best Local Similarity 34.3%; Pred. No. 2.1e-95; Length 1170;
Matches 403; Conservative 210; Mismatches 460; Indels 101; Gaps 37;

QY 1 ENLDTENAMTFO--ENARFGQSVVLOGSRVVGAPQEIIVAAORGSLYQCDYSTGSC 58
Db 26 YMLDVGARSFPFPAGRHFGYRVLOV-GNGVIVGAPGE---GNSGSLYQCSGTGHL 81
QY 59 PIRLQVPEAVNMSLGLSAAATSPQLLACQPTVHQTCTSENTRYKGLCFGLSGMLR--- 115
Db 82 PVTLR-GSNVTSKYLGMWLTATDPTDGSILACDPGLSRTCDQNTYLSGLCYLFRQNLQGP 140
QY 116 -QOPKQFPALRGCPQEDSDIAFLDGSIIIPDPRMKWVSTWMEOLKSKTLFSLM 174
Db 141 LQRPQFCIKG----NVDVFLPFGMSLQDFDFQKILDFKQVVKLSNTISYQFAV 196
QY 175 QYSEEFRIHFTKEFQNNPRSLXPITQLLGRHTATGLRKVVRFLNITNGARONAF 234
Db 197 QFSTYKTEFDGSDYVKRDPDALLKHVKMLLNTTFGAINVATEVPRBELGARPDAT 256
QY 235 KILFLITDEKFGDPLGYEDVPELDRGVIRYVIGVGDAPRSEKSRQRLNTVASKPPRD 294
Db 257 KVLIIITDGE--ATDSGNIDAADK-----IIRYIIGIKGHFOTKESQETLHKFASKP 309
QY 295 HVPQINNFEALTKIQNLKEKIPAEIBGTQTGSSSSPEHEMSQEGFSAATISNGPLISTVG 354

```

RESULT 6

ITAL_MOUSE

```

ID ITAL_MOUSE STANDARD; PRT; 1163 AA.
AC P24063;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-L precursor (Leukocyte adhesion glycoprotein LFA-1
DE alpha chain) (Leukocyte function associated molecule 1, alpha chain)
DE (CD11a).
GN ITGAL OR LFA-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

Db 310 FVKILDTFEKLOLFTLQKIVVIEGTSKQDLTSFNNELSSSGISADLSRGHVAWGAVG 369
QY 355 SYDWAGGVF-LYTSKEKSTFINNTRVDSMDNDAYLGYAAA-IIENRNVQSILVAGAPYQH 412
Db 370 AKDWAGGFLDKADLQDDTFIGNEPLTPEVRAGLYGTVTWLPSRKQKTSLLASGAPYQH 429
QY 413 IGLVAMFR--ONTGWESNANVKTGTGAFGLASLCSVDVDSNGSTDLVLIGAPHYEQT 470
Db 430 MGRVLLPQEPQGGHWSQVQTIHGTYGSGYFGBELGVDDVDQGETELLGLGALFLYGEQ 489
QY 471 RGGQVSVCLPRGORARWQCDAV--LYGEOQPMGRFGAALTVLGDVNGDKLTDAVIGAP 528
Db 490 RGRVFIY-----QRRQLGFEVSELOQDFGYPGRFGAALTALTDINGDLVDVAGAP 544
QY 529 GEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKUSPLRLOYFGQSLSGQDLTMDGLVLT 588
Db 545 LEE--QGAVYIFNGRHG-GLSPQSPQRIEGTVLSSGIGQWFGRSIHGVKDLGEGDLAVAV 601
QY 589 GAQGHVLLLSQPVLRVKATMEFNPREVARNVPCNDQVW-KGKEAGEVRVCLLVQKSTR 647
Db 602 GABSQMIVLSRPPVDMVTLMSPSPAI PVHEVCSTYSNOMKEGVNITICFOI-KSLY 660
QY 648 DRLEEGIQSVVTVYDLALDSGRPHSRAVFNEMTKNSTROTQVLGLTQTCETLKLQLPNCI 707
Db 661 PQF-QGRLVANLYTQLDGHRTRRRGLFPGGRHELNRNIAVT-TSMSCITDFSHPVVCV 718
QY 708 EDPVSPVILRLNPSL---VGTPLS--AFGN-----LEPVLAEQAQLFTALFPKKNCGN 757
Db 719 QDLSPINVSINLPSLWEEBGTFRDQRAQGGKIDIPILRPSLHSETWEI-----PEKNCGE 773
QY 758 DNICQDDLSTTFSPMSLDCLVWGGPREFNVTVVRNDGDSYRTQVTFPPPLDLSYRKVS 817
Db 774 DKCEANLRVSFSPARSALRLTAPASLSVELSLNLEBDAYWVQLDLHPPGLSFRKVE 833
QY 818 TLQNRQSRWRPLACES--ASSTEVSGALKSTSCINHPPEPENSEVTFNFTFDVDSKAS 875
Db 834 ML---KPHSQIPVSCBELPESRLLSAL---SCNWSPIPKAGHSVALQMMFNTLVNS 887
QY 876 LGNKLLKANVTSENN---MPTNKTEFQLELPVKYAVVMVTVSHGVSTKYLNPFTASEN 931
Db 888 WEDSVLEHANTCNEDSDLEDNSATTI---IPLYPINILIQDESDTLVVSFTPKGP 944
QY 932 TSVVWQHOYQV---SNLGQRSLP-ISVFLVPRVRLNQTIVDRPQVTFSENLSTCHTK- 986
Db 945 KIHQVKMYQVRIQPSIHDHNIPTLEAVVGPQPESEGPITHQWSVQMEPPV--PCHYED 1002
QY 987 -ERLPSHSD--FLAEERKAPVWNCNIAVCQRIQCTDIPFGIOEBFNATLKNLSFDWYIK 1043
Db 1003 LERLPDRAAPCLPGALFCPVV-----PRQILVQVIGTLELVGEIE 1044
QY 1044 TSHNELIIVSTABILFNDVSFTLLPGQGAFFRSQTKVEPFPVNPFLIVGSSVGGLL 1103
Db 1045 AS-SMPSLCSLSISFNSSKHFLVYGSNASL-AQVMKVDVVYKQMLYLVLSGIGGL 1102
QY 1104 LLALITAILYKLKFKQYQKMMSEG-GPPGAEP 1136
Db 1103 LLLLIPTVLYKVGPFKRLKKEKMEAGRGVNGIP 1136

```


OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBi_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91268576; PubMed=2051027;
 RA Kaufmann Y., Tseng E., Springer T.A.;
 RT "Cloning of the murine lymphocyte function-associated molecule-1
 alpha-subunit and its expression in COS cells.";
 RL J. Immunol. 147:369-374(1991).
 RN [2]
 RP SEQUENCE OF 24-42.
 RX MEDLINE=85188276; PubMed=3887182;
 RA Springer T.A., Teplow D.B., Dreyer W.J.;
 RT "Sequence homology of the LFA-1 and Mac-1 leukocyte adhesion
 glycoproteins and unexpected relation to leukocyte interferon.";
 RL Nature 314:540-542(1985).
 CC -!- FUNCTION: INTEGRIN ALPHA-L/BETA-2 IS A RECEPTOR FOR ICAM1, ICAM2,
 ICAM3 AND ICAM4. IS INVOLVED IN A VARIETY OF IMMUNE PHENOMENA
 INCLUDING LEUKOCYTE-ENDOTHELIAL CELL INTERACTION, CYTOTOXIC T-CELL
 MEDIATED KILLING, AND ANTIBODY DEPENDENT KILLING BY GRANULOCYTES
 AND MONOCYTES. MICE EXPRESSING A NULL MUTATION OF THE ALPHA-L
 SUBUNIT GENE DEMONSTRATE IMPAIRED TUMOR REJECTION AND IMPAIRED
 LEUKOCYTES RECRUITMENT.
 CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-L
 ASSOCIATES WITH BETA-2.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- TISSUE SPECIFICITY: LEUKOCYTES.
 CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
 WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
 CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
 CC -!- SIMILARITY: Contains 1 VWFA domain.
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 DR EMBL; M60778; AAA39426.1; -.
 DR PIR; I56126; I56126.
 DR HSSP; P20701; ILFA.
 DR MGD; MGI:96606; Itgal.
 DR InterPro; IPR000413; Integrin_alpha.
 DR InterPro; IPR002035; VWF_A.
 DR Pfam; PF01839; FG-GAP; 3.
 DR Pfam; PF00357; integrin_A; 1.
 DR Pfam; PF00092; vwa; 1.
 DR PRINTS; PR01185; INTEGRINA.
 DR PRINTS; PR00453; VWFADOMAIN.
 DR SMART; SM00191; Int_alpha; 5.
 DR SMART; SM00327; WVA; 1.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 DR PROSITE; PS02234; VWFA; 1.
 DR KEGG; Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 KW Signal; Magnesium; Calcium;
 KW Repeat.
 FT SIGNAL 1 23
 FT CHAIN 24 1163
 FT DOMAIN 24 1084
 FT TRANSMEM 1085 1108
 FT DOMAIN 1109 1163
 FT REPEAT 39 88
 FT REPEAT 148 334
 FT DOMAIN 148 334
 FT REPEAT 399 454
 FT REPEAT 455 514
 FT REPEAT 516 573
 FT REPEAT 576 628
 FT CA_BIND 466 474

CA_BIND 528 536
 CA_BIND 588 596
 SITE 1111 1115
 DISULFID 70 77
 DISULFID 108 126
 DISULFID 147 139
 DISULFID 651 705
 DISULFID 767 773
 DISULFID 840 856
 DISULFID 993 1009
 DISULFID 1017 1048
 CARBOHYD 86 86
 CARBOHYD 185 185
 CARBOHYD 270 270
 CARBOHYD 444 444
 CARBOHYD 668 668
 CARBOHYD 696 696
 CARBOHYD 724 724
 CARBOHYD 728 728
 CARBOHYD 776 776
 CARBOHYD 857 857
 CARBOHYD 880 880
 CARBOHYD 899 899
 CARBOHYD 927 927
 CARBOHYD 1056 1056
 SEQUENCE 1163 1163
 A7A30784898232F CRC64;
 Query Match 26.0%; Score 1527.5; DB 1; Length 1163;
 Best Local Similarity 34.1%; Pred. No. 4.5e-94;
 Matches 398; Conservative 217; Mismatches 463; Indels 89; Gaps 36;
 QY 1 FNLDTENAMTQENA-RGFGQSVVQLQGRVVGVAPOEIVAAVQGRGSLYQCDYDYGSCPEP 59
 Db 24 YNLDRPTQSFIAQGRFGVQLQEDG-VVVGARGE---GDNTGGLYHCRSTSEFCQP 79
 QY 60 RLQVVEAVNMISGLSLAATTSPPOLLACQPTVHTQCTSENYVKGFLFGSNLRQOPQ 119
 Db 80 VSLH-GSNHTSKYLQMTLATDAKGLIACDPLGSLRTCDQNTYLSGLCYLPQSLGPM 138
 QY 120 KPEALRCPCPDSPLAFLIDSGSIIIPHDFRMKEWSTVMEQKSKTKTFLSLMQYSHE 179
 Db 139 QNRPAQCEMKGVLDLVELFDGSQLDRKDEKILEPKMVKLSNTSYQFAAVQFSTD 198
 QY 180 FRIHFTFKFP-QNNPNPSLIKPIPOLGRTHATGLKXVRELPNINANGARNKILF 238
 Db 199 CRTEFTLDYVQKNKPNVLLGSGVQPMELLNTFRAINYVVAHVFKBSGARPDATKVLV 258
 QY 239 LIITDGEKF--GDPLGYEDVIPELDREGVIRYVGVGDVAFRSEKSRQELNTVASKPRDHV 296
 Db 259 LIITDGEASDKGNISAHND-----ITRIIGIKKHFVSVOKOKTHIFASEPVEFV 309
 QY 297 FQINNFEALKTIONOLREKIFAIECTQTGSSSPHEHMSQEGFSAAITSNGLLSTVGSY 356
 Db 310 KILDTFEKLDLFTDLQRRIVAIEGTNRQDLTSFNMELSSGSIADLSKGHAVVGAUGAK 369
 QY 357 DWAGGVF-LYTSKESKSTINWTRVDSMDNDAYLGAAR-IILNRVQSVILGAPVQHG 414
 Db 370 DWAGGFLDLREDLQATFVGQEPDLSVDRGGLTVGVAMTSSRSRPLAAGAPRYQHV 429
 QY 415 LVAMER--QNTGMWESNANVKTQTQAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTRG 472
 Db 430 QVLLFQAPFAGGRVWQTKIBGTQIGSYFGGELCSVDLDQGEAEILLIGAPLFGQRG 489
 QY 473 GQVSVCPILPROQARWQCDVLYGBOGQFPGWFRGALTVLGDVNGDKLTDVAIGAPGED 532
 Db 490 GRVFTY--QRRQSLFEMVSELQSDPGYPLGRFGAAITALTIDINGDRLTDVAVGAPLSE- 545
 QY 533 NEGAVYLFHGTSGSGISPSHSORLACSKLSPLOVFGQSLSGQDLTMDGLVDLTVGAGQ 592
 Db 546 -QGAVYIFNKGFG-GLSQPQSGRIQGAQVFFGIRWFGSRHGVKDLGGDLADVVVGA 603
 QY 593 HVLLRSQFVLRVKAIMBFPNPREVARNPECNQDVVKKEAG-FVRVCLHVQKSTEDRLR 651

FT CARBOHYD 1013 1013 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1055 1055 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1086 1086 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1167 AA; 128983 MW; B8331C115DCCCFD CRC64;
Query Match 19.5%; Score 1148.5; DB 1; Length 1167;
Best Local Similarity 28.8%; Pred. No. 116-68;
Matches 355; Conservative 214; Mismatches 459; Indels 205; Gaps 43;
QY 1 FMDTENA--MTQENARGFGSVVQLOGSVVVGAPQBIIVANORGS-----LYQCDY 52
DB ENKVDVWAVTALQCAPAVLSLLHLPN-----NQTCLVARESSNRENTAALYRCAL 74
QY 53 SFGSCPIRLQVPEAVNNSGLSLAATT--SPQILLAC-GPTVHCTCSNTYVXGLCEL 109
DB 75 SI-SPDEIAQ-PVEHICMPKRYQGVTLVNGHGVLCVQVQARKFRSLNSELTGACSL 132
QY 110 FGSNLRQKQKPFALRG-----C-----POB 131
DB 133 LTFNLOQAQYFSDLEGFLDPCGAHVSDGYCRSKGSGTGEEKSARRRRTVEEEDDED 192
QY 132 DSDIAFLIDGSGIIPHDPRMKWVSTWVQL--KSKTILFSLQVSEFRIHTFKPF 189
DB 193 GTEIAIVLDGSGIGSDPKAKNFTSMRNPFYKCFECNFALVQYGAIVQTEFDLQBS 252
QY 190 ONNPNRSLIKPTQLGRTHATGLRKVVRLEPNTNGARKNAFKILLLTGDGKFGDP 249
DB 253 RDINASLAKVQSVQVKEVTKASAMQVLDNIFIPSRGRKALXVWVLDGDIQDP 312
QY 250 LQYEDVPELDREGVYRVYGVGDAPRSEKSEQLNTVASKPRDHVFOINNFEALKTQ 309
DB 313 LNLTTVINSFMQGVVFAVGDRPKNNNTVRELXLIASDPKRAHTPKVWYSALDGLL 372
QY 310 NQREKIPALGQTGSSSFHEMSQGFSAITNSGP--LLSTGVSDWAGSVFLY--TS 367
DB 373 SKLQQRIVEMEGT---VEDALQQLQATGFSQAQLDKQVVLGTGAFNWSGALLYSTQ 429
QY 368 KEKSTFINNT-RVDS-DMNDAYLGYAAAILNRVQSLVGLAPRYQHIGLVAMFRONTQM 425
DB 430 NRGCFNLQNTAKEDSRVTQVSYLYGSLAVLHKAHGISYVAGAPRHKLRGAVFELKRED- 488
QY 426 WESNA---NVKFTQIGAYGASCLVDVNSGSLVGLNGHSLVGLAGHYEYOTGGQVCPPLR 482
DB 489 -EEDAFVRRTEGQMSYFGSVLCPVDIMDGTDFLLVAAPYHIRGEGRYVYQVPE 547
QY 483 GQBARWQCDVLYGEOQPMGRFGAALTIVLGVNGDKLADVAIGAP-----GEEDNRGA 536
DB 548 -QDASFLAHTLSHFGLTNSRFGFAMAAGVDINQKFTDVAIGAFLGFGAGDGASVGS 606
QY 537 VYLFHTSGSGISPSHSORITAGSLSPLOYFQOSLGGQDLTMDGLVLTGQAQCHVLL 596
DB 607 VYTYNGHSG-GLYDPSQIRASSVASGLHYFGMSVSGGLDFNGDGLADITVGSRSADV 665
QY 597 LRQPVLRVKAIMEFNPVARNVFECDQVVGKEAGEVRVCLHVOKS---TRDLRREG 653
DB 666 LRARPVDLTVSMTFP-----DALPMVPIGKM--DVNLCFEYDSSVASEPGLREM 715
QY 654 QIOSVVTYDALDSGRPHSAVNETNSTRTOVLGLTQTC-----696
DB 716 FLNFTVDVW-----TKQRLQCEDSSGQSCLRKWNKSGSFLCEHFWLI 760
QY 697 ETLLQLPNCIEDPVSPIVLRNLFSLVGTSLAPGNLR-----PVLAEADQLFTALF--P 750
DB 761 STEEL-----CEDCFSPNITIKYVE-----PQTSGGRRDPNPTL--DHVKPSALFQLP 809
QY 751 FEKNCGNDNTCQDLSITFSMGLDLVWGSPREFNVTVTRVNDGSDSYRTQVTFPPDL 810
DB 810 YEKDCNKNVFCIAEIQLTN--ISQELVGVGVTKTEVMTNISLTNSGSDSYMTNMLNYPN 868
QY 811 LSYRKVSTLQNSORSRWLACASSTEVSGALKSTCSINHPIPFENSEVTNFTPDV 870
DB 869 LQFKKI-----QKPSVDPQCCDDPFX--ASVLWNKIGHPIL-KRSSVNVSVTWQL 918

QY 871 DSKASLGNKLLKANKYTSNNMPTNKTEPQLELPVKYAVVWVTSCHVSKYLNFTASE 930
DB 919 EESVFPNRTADITVTSNSENKSLARETR--SLQFRHAFIAVLRS--PSVWYMN--TSQ 971
QY 931 NTSRVMOHOYVSNLQORSLSPLSVFLVPLVRLNQTVIWDPRQVTFSENLST-----CHT 985
DB 972 SPDSHKPEFFNVHGENLFGAVFQIQICVPIKQDF-----QIVRVKNLTKQDHTECTQ 1025
QY 986 KERLPSSHSDFLAELRKAPVNVCSIAVCQRIQCDIPFGIQEEFNATLKNLSFDWYIKTS 1045
DB 1026 SQEPACGSDPQVQHKWESVVCAL-----TSNKENYTVAAEISVVG 1065
QY 1046 HNELLVISTA-----EILFNDVSFTLLPGCGAFVRSOTETKVBPF-----EVPNPLPLIV 1095
DB 1066 HTKQLLRVDSLEPILGEISFNKSLYEGNAE-----NHRKTIIVFLKBEETSLPLII 1119
QY 1096 GSSVGGILLALLITALYALYKLGFFKRYKOMMSE 1128
DB 1120 GSSIGGLVAVVIIAILFKCGFFKRYQOQLNLE 1152

RESULT 8

ITAE HUMAN
ID ITAE HUMAN STANDARD; PRT; 1179 AA.
AC P38570; Q9NZ9;
DT 01-OCT-1994 (Rel. 30, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-E precursor (Mucosal lymphocyte-1 antigen) (HML-1
DE antigen) (CD103 antigen) (Integrin alpha-IBL).
GN ITGAE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini, Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1] TaxID=9606;
RP SEQUENCE FROM N.A., AND SEQUENCE OF 19-38 AND 179-188.
RC TISSUE=Leukemia, and Lymphocytes;
RX MEDLINE=94164962; PubMed=811947;
RA Shaw S.K., Cepek K.L., Murphy E.A., Russell G.J., Brenner M.B.,
RA Parker C.M.;
RT "Molecular cloning of the human mucosal lymphocyte integrin alpha E
RT subunit. Unusual structure and restricted RNA distribution.";
RL J. Biol. Chem. 269:6016-6025 (1994).
RN [2]
RP REVISIONS TO 88-114.
RA Parker C.M.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 53-1179 FROM N.A.
RC TISSUE=Fetal kidney;
RX MEDLINE=20138496; PubMed=10673275;
RA Touchman J.W., Anikster Y., Dietrich N.L., Maduro V.V., McDowell G.,
RA Shetlersuk V., Bouffard G.G., Beckstrom-Sternberg S.M., Gahl W.A.,
RA Green E.D.;
RT "The genomic region encompassing the nephropathic cystinosis gene
RT (CTNS): complete sequencing of a 200-kb segment and discovery of a
RT novel gene within the common cystinosis-causing deletion.";
RL Genome Res. 10:165-173 (2000).
RN [4]
RP MUTAGENESIS OF ASP-109 AND PHE-316.
RX MEDLINE=20400502; PubMed=10837471;
RA Higgins J.M.H., Cernadas M., Tan K., Irie A., Wang J.-H., Takada Y.,
RA Brenner M.B.;
RT "The role of alpha and beta chains in ligand recognition by beta 7
RT integrins.";
RL J. Biol. Chem. 275:25652-25664 (2000).
CC -!- FUNCTION: INTEGRIN ALPHA-E/BETA-7 IS A RECEPTOR FOR E-CADHERIN. IT
CC MEDIATES ADHESION OF INTRA-EPITHELIAL T-LYMPHOCYTES TO EPITHELIAL
CC CELL MONOLAYERS.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA
CC SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAINS LINKED BY A
CC DISULFIDE BOND. ALPHA-E ASSOCIATES WITH BETA-7.

Db 757 SGLQCELDL-LMPTEGELCEEDCFNSASVKSQYL-QTPGQTDHPQIPILORYTEPPAI 814
Qy 747 ALPFFPKNGNDNICODDISITFSFMSLCLVVGGRPREFNVTVVRNDGDSYRIQVTFP 806
Db 815 FQLPYEKACKNKLFCVAELQLA-TTVSQQLVWGLTKELTLNLTNSGDSYMTSMALN 873
Qy 807 FPLDLSVRKYSTIONORSQSWELACESASTVSGALKSTCSINHPIFPENSEVTFNI 866
Db 874 YPNLQ-----LKRQKPPSPNIQDDPPV---ASVLMNCRIGHVPL-KRSSAHVSV 923
Qy 867 TFDVDSKASLGNKLLKANVTSSN-----MFRNKTEFO---LELPVKVAVYVVTSHGV 919
Db 924 VQOLENAPNRTADITVTNSNERRSLANETHILQFRHGFVAVLSKPSIMYVNTGQGL 983
Qy 920 S--TKYLNFTASENTSVMOHQOVSNLQORSLSPLSLVPLVVRMLQTVINDRPOVTFSE 977
Db 984 SHKEFLFHVHGEN---LFGAEYQ-----LQICVTKLRLGLQVAAVKLTTRQ 1028
Qy 978 NLSSTCHTKERLPSHSDPLAELRKAPVYVNCISAVCQRIQDIPFGIQBEFNATLKNLS 1037
Db 1029 ASVTQWSQERACAYSS-VQVVEHMSVSCVIA-----SOKENVTVAAIS 1073
Qy 1038 PDWYKTSNHLILVST-----AEILFNDVSFTLLPGQAFVRSQTEIKVEFEVFNPL 1091
Db 1074 WD-----HSEBLKDVTELOILGEISFNKSLYEGLNENH--RTKITVTVFLKDEKXHS 1125
Qy 1092 PLIVSGSVGGLLILALITAAALYKLGFPKQYKDMSE 1128
Db 1126 PIIKSGVGLLVILVILVILKPGFPKRYQQLNLB 1162

RESULT 9
ID _ITAL HUMAN STANDARD; PRT; 1151 AA.
AC P56199;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Integrin alpha-1 (laminin and collagen receptor) (VLA-1) (CD49a).
GN ITGAL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=93155124; PubMed=8428973;
RA Briesewitz R., Epstein M.R., Marcantonio E.E.;
RT "Expression of native and truncated forms of the human integrin alpha
1 subunit."
RL J. Biol. Chem. 268:2989-2996(1993).
CC -!- FUNCTION: INTEGRIN ALPHA-1/BETA-1 IS A RECEPTOR FOR LAMININ AND
COLLAGEN. IT RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-P-P-G-
E-R IN COLLAGEN.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-1
ASSOCIATES WITH BETA-1.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -!- DATABASE: NAME=PRO; NOTE=CD guide CD49a entry;
WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd49a.htm".
CC PIR; A45226; A45226
CC PDB; 1QC5; 17-MAY-00.
CC Genew; HGNC:6134; ITGAL.
CC MIM; 192968; -
CC GO; GO:0008305; C:integrin complex; TAS.
CC GO; GO:0004895; P:cell adhesion receptor activity; NAS.
CC GO; GO:0005518; P:collagen binding; TAS.
CC GO; GO:0007160; P:cell-matrix adhesion; NAS.
CC InterPro; IPR000413; Integrin_alpha.

DR InterPro; IPR002035; VMF A.
DR Pfam; PF01839; FG-GAP; 3-
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR SMART; SM00191; Inc_alpha; 5.
DR SMART; SM00327; VWFA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS02334; VWFA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
Repeat; Calcium; Magnesium; 3D-structure.
FT DOMAIN 1 1113 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1114 1136 POTENTIAL.
FT DOMAIN 1137 1151 CYTOPLASMIC (POTENTIAL).
FT REPEAT 16 75 FG-GAP 1.
FT REPEAT ? ? FG-GAP 2.
FT REPEAT ? ? FG-GAP 3.
FT REPEAT 147 360 VWFA.
FT REPEAT 349 404 FG-GAP 3.
FT REPEAT 405 457 FG-GAP 4.
FT REPEAT 459 520 FG-GAP 5.
FT REPEAT 540 599 FG-GAP 6.
FT REPEAT 602 654 FG-GAP 7.
FT CA_BIND 470 478 POTENTIAL.
FT CA_BIND 552 560 POTENTIAL.
FT CA_BIND 614 622 POTENTIAL.
FT SITE 1119 1142 GPFKR MOTIF.
FT DISULFID 54 64 BY SIMILARITY.
FT DISULFID 660 669 BY SIMILARITY.
FT DISULFID 675 728 BY SIMILARITY.
FT DISULFID 780 786 BY SIMILARITY.
FT DISULFID 850 858 BY SIMILARITY.
FT DISULFID 1002 1034 BY SIMILARITY.
FT DISULFID 1037 1044 BY SIMILARITY.
FT CARBOHYD 46 46 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 72 72 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 77 77 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 84 84 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 189 189 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 289 289 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 313 313 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 374 374 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 432 432 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 504 504 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 671 671 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 720 720 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 752 752 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 812 812 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 855 855 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 880 880 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 887 887 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 911 911 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 938 938 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 946 946 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 980 980 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1045 1045 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1055 1055 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1074 1074 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1085 1085 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1151 AA; 127837 MW; 6B3F3C1AABF52808 CRC64;

Query Match 18.6%; Score 1093.5; DB 1; Length 1151;
Best Local Similarity 27.8%; Pred. No. 5.2e-65;
Matches 345; Conservative 206; Mismatches 487; Indels 203; Gaps 44;

Qy 1 FNLDTENAMTFQENARG-FGQSVVQL-----QGRVVVVGAFQEIIVANQGRSLVQCDYSTGS 55
Db 1 FNVDVKNMTFSGPVEDMFGYTVQYENEEGKWLIGSLVQGPKNRTGDVYKCPVGRGE 60
Qy 57 CEP-IRLQVPVEA-----VNMISLGLSLAATTPSPQLACGPVTHQTCSENTYVKG 106
Db 61 SLEPCVKLDLPVNTSINPVTEVKENMTFGSTL-VTNPGGFLACGPLYAROGHLHYTTGI 119
Qy 107 CFLFGSNLRQQPQKPFPEALRGCTQEDSDIAFLDGSISIPHDFRKNKENVSTVMSQLKK 166

```

DB      1111 RVPLWVILLSAPAGLLMLLHLALWIKIGFFPKRPLKKQWEK 1151
RESULT 10
ITAH_HUMAN
ID      ITAH_HUMAN          STANDARD;          PRT; 1189 AA.
AC      Q9UKX5; Q9UKQ1;
DT      16-OCT-2001 (rel. 40, Created)
DD      16-OCT-2001 (rel. 40, last sequence update)
DT      28-FEB-2003 (rel. 41, last annotation update)
DE      Integrin alpha-11 precursor.
GN      ITG11.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RS      SEQUENCE FROM N.A.
RT      TISSUE=Petal heart, and Osteoblast;
RX      MEDLINE=994117678; PubMed=10486209;
RA      Lehnert K., Ni J., Leung E., Gough S.M., Weaver A., Yao W.P., Liu D.,
RA      Wang S.-X., Morris C.M., Kissansen G.W.;
RT      "Cloning, sequence analysis, and chromosomal localization of the novel
RT      human integrin alphai1 subunit (ITG11).";
RL      Genomics 60:179-187(1999).
RN      [2]
RS      SEQUENCE FROM N.A.
RT      TISSUE=Petal muscle, and Uterus;
RX      MEDLINE=99395147; PubMed=10464311;
RA      Velling T., Kusche-Gullberg M., Seijersen T., Gullberg D.;
RT      "cDNA Cloning and Chromosomal Localization of Human Alpha(11)
RT      Integrin. A collagen-binding, i domain-containing, beta(1)-associated
RT      integrin alpha-chain present in muscle tissues.";
RL      J. Biol. Chem. 274:25735-25742(1999).
RN      [3]
RS      SEQUENCE OF 954-1188 FROM N.A.
RT      TISSUE=Fibroblast;
RX      Andreu N., Estivill X., Escarceller M., Sumoy L.;
RT      Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RL      1-!- FUNCTION: INTEGRIN ALPHA-11/BETA-1 IS A RECEPTOR FOR COLLAGEN.
RL      2-!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-11
RL      ASSOCIATES WITH BETA-1.
RL      3-!- SUBCELLULAR LOCATION: Type I membrane protein.
RL      4-!- TISSUE SPECIFICITY: ACCORDING REF.2 HIGHEST LEVELS IN UTERUS AND
RL      HEART, INTERMEDIATE LEVELS IN SKELETAL MUSCLE AND INTERMEDIATE TO
RL      LOW LEVELS IN PANCREAS, KIDNEY AND PLACENTA. ACCORDING REF.1 ALSO
RL      FOUND IN BRAIN, COLON, LUNG, SMALL INTESTINE, STOMACH, TESTIS,
RL      SALIVARY GLANDS, THYROID GLANDS AND PROSTATE. VERY LOW LEVELS IN
RL      PERIPHERAL BLOOD LYMPHOCYTES, FETAL BRAIN AND FETAL LIVER.
RL      5-!- DEVELOPMENTAL STAGE: STRONGLY UP-REGULATED IN DIFFERENTIATING
RL      FETAL MUSCLE CELLS (IN VITRO).
RL      6-!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
RL      WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
RL      7-!- SIMILARITY: Belongs to the integrin alpha chain family.
RL      8-!- SIMILARITY: Contains 1 VWFA domain.
RL      9-!- SIMILARITY: Contains 7 FG-GAP repeats.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
EMBL; AF109681; AF01258.1; -
EMBL; AF137378; XAD51919.2; -
EMBL; AL359064; CAB94392.1; -
HSSP; P17301; LAGX
HGNC; HGNC:6136; ITG11.
MIM; 604789; -
GO; GO:0008305; C:integrin complex; TAS.
GO; GO:0004895; F:cell adhesion receptor activity; TAS.

```


1048 PVE--EDLRRAPQLNHSNSDVVSINCNIRLVP-NQEIFHLLGNL---WLSRLKALKYKS 1101
 1047 NEHLIVSTAETILFNDVSFTLLPQCGAFVRSQETRVSEFEVFN-----PLPLIVGSSVG 1100
 1102 MKIMVNAALQROFH-SPF-----IFREDEPSQIVFBISKQBEDWQVPIIIVGSLTG 1152
 1101 GLLALALTAALYKLGFTK--ROYKDMMSBEGPPQABP 1136
 1153 GLLALALVLALWKLGFERSARRRE-----PGLDLP 1183

RESULT 11
 ITA2_BOVIN STANDARD; PRT; 1170 AA.
 AC PS3710;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Integrin alpha-2 precursor (Platelet membrane glycoprotein Ia) (GP1a)
 DE (Collagen receptor) (VLA-2 alpha chain) (CD49b) (Fragment).
 GN ITGA2.
 OS Bos taurus (Bovine).
 OC Sukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OC NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94193647; PubMed=75111592;
 RA Kamata T., Puzon W., Takada Y.;
 RT "Identification of putative ligand binding sites within I domain of
 integrin alpha 2 beta 1 (VLA-2, CD49b/CD29).";
 RL J. Biol. Chem. 269:9659-9663 (1994).
 CC -!- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A RECEPTOR FOR LAMININ,
 COLLAGEN, COLLAGEN C-PROPEPTIDES, FIBRONECTIN AND E-CADHERIN. IT
 CC RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-F-P-G-E-R IN
 CC COLLAGEN. IT IS RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER
 CC CELLS TO COLLAGENS. MODULATION OF COLLAGEN AND COLLAGENASE GENE
 CC EXPRESSION, FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED
 CC EXTRACELLULAR MATRIX.
 CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-2
 CC ASSOCIATES WITH BETA-1.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
 CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
 CC -!- SIMILARITY: Contains 1 VWFA domain.
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; L25886; AAB59255.1; -
 CC PIR; I45914; I45914.
 CC HSSP; P17301; LAOX.
 CC InterPro; IPR000413; Integrin_alpha.
 CC InterPro; IPR002035; VWF_A.
 CC Pfam; PF01839; FG-GAP; 3.
 CC Pfam; PF00357; Integrin_A; 1.
 CC Pfam; PF00092; vwa; 1.
 CC SMART; SM00191; Int_alpha; 5.
 CC SMART; SM00327; VWA_1
 CC PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 CC PROSITE; PS00234; VWFA; 1.
 CC Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 CC Platelet; Signal; Repeat; Polymorphism; Calcium; Magnesium.
 CC NON_TER 1 1

SIGNAL <1 18
 CHAIN 19 1170
 DOMAIN 19 1121
 TRANSMEM 1122 1143
 DOMAIN 1144 1170
 REPEAT 34 92
 REPEAT ?
 REPEAT 177 367
 DOMAIN 423 475
 REPEAT 477 538
 REPEAT 540 599
 REPEAT 604 656
 CA_BIND 488 496
 CA_BIND 552 560
 SITE 472 474
 SITE 1146 1150
 DISULFID 72 81
 DISULFID 669 726
 DISULFID 778 784
 DISULFID 854 865
 DISULFID 1008 1039
 DISULFID 1044 1049
 CARBOHYD 94 94
 CARBOHYD 101 101
 CARBOHYD 322 332
 CARBOHYD 421 421
 CARBOHYD 449 449
 CARBOHYD 464 464
 CARBOHYD 688 688
 CARBOHYD 748 748
 CARBOHYD 945 945
 CARBOHYD 1063 1063
 CARBOHYD 1070 1070
 VARIANT 580 580
 VARIANT 588 588
 VARIANT 725 725
 SEQUENCE 1170 AA; 138929 NM; ECEFC1C5F2448FB1 CRC64;

Query Match 18.2%; Score 1071; DB 1; Length 1170;
 Best Local Similarity 27.4%; Pred No. 1.7e-63;
 Matches 333; Conservative 219; Mismatches 435; Indels 168; Gaps 47;

QY 1 FNLDENAMTFQ-ENARGFGOSVQL---QSSRVVVGAPQEIIVAAVNGSGSYQC--DYST 54
 19 YNVLPRKAFIPSGPSSEQFGYAVQOFINPKGNWLLVGSFPMSPGPKNRMGDVYKCPVDLST 78
 55 GSCPIRLQ-----VPVEAVNMSLGLSLAATTSPOLLACGPTVHCTCSENVVVKGLC 107
 79 TTCEKLNLTSTMSNVTEMKTNMSLGLTFRNVGTGGFLTCGLPMAQQCCSQYTTTGV 138
 108 PLFGSNLRQPKFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMRKEMVSTVMEQLK-- 165
 139 SDVSPDF-QLRTSPAPAVQTCF-SFIDVVVVCDESNSIYPWD--AVKNLEKFKVQGLDIG 194
 166 KSKTLFSLMQVSEBFRHFTFKFQNNPNRSLIKPTOLL----GRTHATGLKRVRE 221
 195 PTKQMGILQVANNRPNVFNLTFSKD---EMIKATSTQFYQGGDLTNTFKAIQADRT 251
 222 LFNITNGARKNAFKILPLLTDGEKFGDPLGVEDYIPELDRGVIRYVIGV-----GDAPR 276
 252 AYSTAAGRPGATKVMVVTDGESH-DGSKLKAVIDQCNKILRFGIAGVLYLNALD 310
 277 SEKSRQELNTVASKPPRDHVFQINNFALKTIONOLREKIPAIKRGTTQGTSSSSEHMSQ 336
 311 TKNLKEIKALIASIPTEHPFNVSDEADLLEKAGTIGEQIFSIETVQG-GDNFQMEMSQ 369
 337 EGFSNAIT--SNGPLISTVGSYDWDAGVFLVTSKESKINMT--RVDSDMN-DAYLGYA 391
 370 VGFSAEYSPQNNILMLGAVGAYDSGTVQKTPHGLIFSQAPEQIILQDRNHSSLYGS 429
 392 AAIIILNRVQSLVLGAPRYQHIGLVAMPFRONTGMWESNANV-----KGTQIGAYFGASL 445


```
Db 430 VASISTGNSVHFVAGAPRANTTCQIVLYSVN-ENGNTVTIQSGDQIGSYFGSVL 484
Qy 446 CSVDVDSNGSTDLVLCAPHYEQTR--GGVSVCPLPRGORARWCCDAVLXGEOQOPWG 503
Db 485 CAVDVNDITDVLVAGPYNMDLKEBGRVFLFTIKG--ILNWH--QFLGPNGLENA 541
Qy 504 RFGAALTVLGVNDGKLTDLVAIGAPEDNRGAVYLFHGTSGSGISPSHSQRIAGS--KL 561
Db 542 RFGSAAALSDINDMGDFNDVIGSPLENQNSGAVTYNGHEG--IRLRSQKILSDRAF 600
Qy 562 SPRLQYFGOSLSGQDITMDGLVLTGAGQHVLLRSOPVLRLVKALMEFNEPREVARNVF 621
Db 601 SSLQYFGRSLDYGDLNGSDITFVSAGPQVQVQVQVQVQVQVQVQVQVQVQVQV 658
Qy 622 ECNDQVVKGBEVRVCLHVQKSTDRRLRREGQIQSVVTVYDLDL-----SGRPHSRAVFN 677
Db 659 NKVAET-----KLXCLP---SARFPTNQNQVAIVNITIDEDQSSRVISRGLEK 707
Qy 678 ETKNSTRQTVGLGTQCE--TLKQLPNCIEDPSPVIRLNPSL--VGTPLSAFGLN 733
Db 708 ENNERCLQKTMIVSQQRSEYIIHQEPS---DIISPLNLCNWSLENPGT-----756
Qy 734 RPYLADAOQLFTALPFENKNCNDNICODLSITF-----SFMSLDCLVVGGRPFNFVTV 789
Db 757 NPALBAYSETVKVPSIPFHKDCGDGVCISDLVNVQQLPATQOQPFIVSNQNKRLTFSV 816
Qy 790 TVRNDGEDSVRTQVTFPFLDLISYRKVSTLONORSORSHLACESASST-EVSGALKSTFS 848
Db 817 QLRNKKESAYNTEIVDFSENLPF-----ASWMPVDGTEVTCQIASSQKSVT 864
Qy 849 CSNHPHIPPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEFOLELPVKY 908
Db 865 CNVGYFALKSKQVTFINEFDNLQ--NLQNASISFRALSESDEENWADSNVNLKSLLY 923
Qy 909 AVTMVTSHGVTKYNLNFASNTSRVMQHYQVSNLQOR-----SLPISLVFLV 958
Db 924 DABIHIT-RSTNINFEVSLDGNVSVV-HSFE--DIGPKFIPSIXVTGTSVPSVNA---976
Qy 959 FVRLNQTIVNDRPOVTFSEN-----LSSCTHCKE-----RLPHSDSLAB- 998
Db 977 -----SVIIHIOYTKDKNPLMYLTVGVHTDQAGDISCEABINPLKIGQTSVSVPSKSEN 1030
Qy 999 LRKAPVNVCSIAVCORIQCDIPFGIQEFNATLKNLSDFDVYIKTSHNHLIVSTAB- 1057
Db 1031 FRHKELNCRKTASCSNIMCMLRDLQVKGFFLVNSTRINGTFAASTFTQVQLTAAAEID 1090
Qy 1058 LFVDSVFTL-----LPQGAFAVRSQTEKVEPFE-VENPLPLIVGSSVGGILLALITA 1110
Db 1091 TYNPQIYVISENTVTIP-----LTIKPKHEKVEVPTGVIVGSIAGILLALLALVA 1140
Qy 1111 ALYKLAGFFKQYKDM 1125
Db 1141 ILWKLGFKKYKEM 1155

RESULT 12
ID ITA2 MOUSE STANDARD; PRT; 1178 AA.
AC Q62469; Q62163;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-2 precursor (platelet membrane glycoprotein Ia) (GP1a)
GN (Collagen receptor) (VLA-2 alpha chain) (CD49b).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6 X CBA; TISSUE=Lung;
```

```
RX MEDLINE=94363406; PubMed=8081889;
RA Edelman J.M., Chan B.M., Uniyal S., Onodera H., Wang D.Z.,
RA Damjanovich L., Latzer D.B., Finberg R.W., Bergelson J.M.;
RT "The mouse VLA-2 homologue supports collagen and laminin adhesion but
RL not virus binding.";
RN Cell Adhes. Commun. 2:131-143(1994).
RP SEQUENCE OF 450-1178 FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=94355691; PubMed=7521231;
RA Wu J.B., Santoro S.A.;
RT "Complex patterns of expression suggest extensive roles for the alpha
RL 2 beta 1 integrin in murine development.";
CC Dev. Dyn. 199:292-314(1994).
CC -!- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A COLLAGEN RECEPTOR, BEING
CC RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER CELLS TO
CC COLLAGENS, MODULATION OF COLLAGEN AND COLLAGENASE GENE EXPRESSION,
CC FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED
CC EXTRACELLULAR MATRIX. IT IS ALSO A RECEPTOR FOR LAMININS, COLLAGEN
CC C-PROPEPTIDES AND E-CADHERIN. MICE HOMozyGous FOR A NULL MUTATION
CC IN THE ALPHA-2 DIE VERY EARLY IN EMBRYOGENESIS.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-2
CC ASSOCIATES WITH BETA-1.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; Z29987; CAA82877.1; -
DR EMBL; X75427; CAA53178.1; -
DR PIR; S44142; S44142.
DR HSP; P17301; LAOX.
DR MGD; MGI:96600; Itga2.
DR InterPro; IPR000413; integrin_alpha.
DR InterPro; IPR002035; VWF A.
DR Pfam; PF01839; FG-GAP 3.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA_1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWFA; 1.
DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Platelet; Signal; Repeat; Calcium; Magnesium.
FT SIGNAL 1 26
FT CHAIN 27 1178
FT DOMAIN 27 1129
FT TRANSMEM 1130 1151
FT DOMAIN 1152 1178
FT REPEAT 42 100
FT REPEAT ? ?
FT DOMAIN 185 375
FT REPEAT ? ?
FT REPEAT 431 483
FT REPEAT 485 546
FT REPEAT 548 607
FT REPEAT 612 664
FT REPEAT 496 504
FT CA_BIND 560 568
FT CA_BIND 624 632
FT SITE 480 482
FT DISULFID 80 89
FT SITE 1154 1158
FT BY SIMILARITY.
FT CELL ATTACHMENT SITE (POTENTIAL).
FT OFFER MOTIF.
FT BY SIMILARITY.
```

Query Match 18.0%; Score 1057; DB 1; Length 1178;
 Best Local Similarity 27.8%; Pred. No. 1.5e-62;
 Matches 340; Conservative 210; Mismatches 487; Indels 188; Gaps 44;

FT DISULFID 677 734 BY SIMILARITY.
 FT DISULFID 786 792 BY SIMILARITY.
 FT DISULFID 862 873 BY SIMILARITY.
 FT DISULFID 1016 1047 BY SIMILARITY.
 FT DISULFID 1052 1057 BY SIMILARITY.
 FT CARBOHYD 102 102 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 109 109 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 429 429 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 457 457 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 472 472 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 696 696 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1054 1054 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1071 1071 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1078 1078 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1178 AA; 128926 MW; 1F194B9C0240F465 CRC64;

QY 1 FNLDTENAMTEQ-ENARFGQSVQL---QGSVVVVGAPQEIIVAAANORGLSYQC---DYST 54
 DB 27 YNVLPGAKIFGSPSEQSGYVQQLTNPQGNLLVGSFSPFENRMGSDVYKCPVDLPT 86
 QY 55 GSCEPIRLQ-----VPVEAVNMISGLSLAATTSPQQLACGPTVHQTCSNTYVKGIC 107
 DB 87 ATCEKLNQNSASISNVTIKTMSLGLTLTNPCTGGFLTCGLPLWAHQCGNQYATGIC 146
 QY 108 FLFGNLRQPOQ---KPPBALRGCTQEDSDIAFLIDGSGIIPHFRKKEWVSTWMEQL 164
 DB 147 ----SDVSPDFQFLTSFSPAVQACPSL-VDVVVVCDSENSIYP--WEAVKFLKVFVTGL 199
 QY 165 X--KSKTFLPSLMQYSEEPRIHPTFKFQNNPNSRLIKPITQLLG-RHTATGLKRVRE 221
 DB 200 DIGPKTKTQVALQYANEPRIIFNLNDFETKEDVQATSTROHGGDLTWTFPAIFARDY 259
 QY 222 LFNITNGARKNAFKLFLTLTGCKFGDPLGYEDVPELDRBGVIRYICV-----GDAPR 276
 DB 260 AYSQTSGRPGATKVMVWVTDGESH-DGSKLKTIVIQCCNDDBILRFGIAGVLYLRNALD 318
 QY 277 SEKSRELNTVASKPRDRHVFQINNEALTKIQNLREKI FALGCTGTGSSSPHEMSQ 336
 DB 319 TKNLKEIKAIKATPTERTFFNVADBALEKAGLTGEQIFSLGTVQG-GDNFQNEHQ 377
 QY 337 EGFS--ALTSNGPLSLTVSGYDAGVFLYTSKEKSTFINMT--RVSDSMN-DAYLGYA 391
 DB 378 VGFSAADYAFQNDILMLGANGAFDWSCTLVQETSKHPVFPKQAFQVLQDRNHSSFLGYS 437
 QY 392 AAILLRNVQSLVGLAPRYQHIGLVAMFRONTGMWESNANV-----KGTQIGAYFGASLCS 447
 DB 438 VAAISTEDGVHFFVAGAPRANYTQGVLYSVNK---QGNVTVIQSHRGDQIGSYFGSVLCS 494
 QY 448 VDVSNGSLDLVLIGAPHYVEQTR--GGQVSVCPPLPRGQARWQCDAVLYGQGQGPWGRF 505
 DB 495 VDVKDTITDVLVGAPTYNDILKEEGKLYLFTITKGLNCHQ---FLEGEGTGNARF 551
 QY 506 GAALTVLGVNDGKLFTVAIGAPEGEDNRGANVLFHGTSGSGISPSHSORIASGLSPR 564
 DB 552 GSAIAALSINMDGFNDVIGSVENSGAVIYNGHQGT-IRTKYSQKILGSGNGAFR 610
 QY 565 -LQYFGOSLGGGDLTMDGLVLTVCAGQGHVLLRSCQPLRVKAIKMFNPREVARNVPEC 623
 DB 611 HQFFGRLDYGDLNGDSITDVSIGALQVLIQWSQSDADVAIEALFTP----- 660
 QY 624 NDQVVKGEAGEVRLVCHVKSTRDLREGQIQSVVYDLDL----SGRPHSRVFNFT 679
 DB 661 -DKITLNLKDAKITLCLFAPRPAQNNQV--AILFNNTLDADGHSSRVTSRGVFRN 717
 QY 680 KNSTRQTVLSTQTCET--LKLQPLNCIEDPVSIVLRLAFESVGTPLSAFGLNRPVL 737
 DB 718 SERFLQNMVNVNQQKSEHHISIQKPS---DVNPLDLRVDLISLENPGTS-----PAL 768
 QY 738 AEDAQLRFTALPPFKKNCNDNICQDLSL-----TFGMSLDCDLVWGSPREFNVT 788

Db 769 EAYSETVKVPSIPFYKEGSDGICISDLILDVQOLPAIQOSP-----IVSNQKRLTFS 823
 QY 789 VYVRNDEGDSYRTQVTFFPFFELDLSYRKVSTLQNRQSRQSWRLACESASST-EVSGALKST 847
 Db 824 VILKNRGESAYNTVLAEPSENLF-----ASPSMPVDGTEVCEVGSQSKSV 871
 QY 848 SCSINHIPIFENSEVNTNITFDVDSKASLGNKLLKXANVTSENNMPRTNKT--FOLELP 905
 Db 872 TCDVGYFALKEQOQVTFITNFDFNLQ-NLQNAINFAQFSESQ--ETNKADNSVSLTIP 928
 QY 906 VKYAVYVAVTSHGVSTKYLNFTASENTSRVMQHOYQVSNLQOR-----SLPISLV 955
 Db 929 LLYDAELHLT-RSTNINFISSDENAPSVIK---SVEDIGPKFIFSLKVTAGSAPVMA 984
 QY 956 PLV-----PVRLNQTVINDRQVTF--SENLS 980
 Db 985 LVTHIPQYTKERNPLLYLTGIDQDAGDISCTABINPLKLPHTA----PSVSPKNENFR 1040
 QY 981 STCHTKERLPSSHDFLAELRKAPVVGSIACVQRIQCDIPFGIQEEFNATLKNLSFDW 1040
 Db 1041 ---HTKE-----LDCRTTSCSNITCKLXDLHMKAEYFINVTTVRVWRT 1080
 QY 1041 YIKTSHNLLIVSTABILEFNDVFTLLPGQAGAFVRSOTETKVFSEFFVNPPLPLVGSVG 1100
 Db 1081 FAASTFTQVLTAAGAEIDTNPQLFVIEENAVTIPLMIMKTEKAEVPT--GVIIIGSIIA 1138
 QY 1101 GLALLALITAAVLYKLGPFKROYKDM 1125
 Db 1139 GILLALLATAGLWKGFFKROYKDM 1163

RESULT 13
 ITAL RAT STANDARD; PRT; 1180 AA.
 AC P16614;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Integrin alpha-1 precursor (Laminin and collagen receptor) (VLA-1)
 DE (CD49a).
 GN ITGAL.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90338125; PubMed=2380249;
 RA Ignatius M.J., Large T.H., Houde M., Tawil J.W., Barton A.,
 RA Esch F., Carbonetto S., Reichardt L.F.;
 RT "Molecular cloning of the rat integrin alpha 1-subunit: a receptor
 RT for laminin and collagen.";
 RL J. Cell Biol. 111:709-720(1990).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 151-364.
 RX MEDLINE=99313197; PubMed=10386626;
 RA Nolte W., Pepinsky R.B., Vennyaminov S.Y., Koteliarsky V.,
 RA Gotsals P.J., Karpusas M.;
 RT "Crystal structure of the alphabeta integrin I-domain: insights into
 RT integrin I-domain function.";
 RL FEBS Lett. 452:379-385(1999).
 CC -!- FUNCTION: INTEGRIN ALPHA-1/BETA-1 IS A RECEPTOR FOR LAMININ AND
 CC COLLAGEN. IT RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-F-P-G-
 CC E-R IN COLLAGEN.
 CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-1
 CC ASSOCIATES WITH BETA-1.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VFMA DOMAIN. INTEGRINS
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
 CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
 CC -!- SIMILARITY: Contains 1 VFMA domain.
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.

FT	DISULFID	83	92	BY SIMILARITY.	
FT	DISULFID	680	737	BY SIMILARITY.	
FT	DISULFID	789	795	BY SIMILARITY.	
FT	DISULFID	865	876	BY SIMILARITY.	
FT	DISULFID	1019	1050	BY SIMILARITY.	
FT	DISULFID	1055	1060	BY SIMILARITY.	
FT	CARBOHYD	105	105	N-LINKED (GLCNAC. . .)	(POTENTIAL) .
FT	CARBOHYD	112	112	N-LINKED (GLCNAC. . .)	(POTENTIAL) .
FT	CARBOHYD	343	343	N-LINKED (GLCNAC. . .)	(POTENTIAL) .
FT	CARBOHYD	432	432	N-LINKED (GLCNAC. . .)	(POTENTIAL) .
FT	CARBOHYD	460	460	N-LINKED (GLCNAC. . .)	(POTENTIAL) .
FT	CARBOHYD	475	475	N-LINKED (GLCNAC. . .)	(POTENTIAL) .
FT	CARBOHYD	699	699	N-LINKED (GLCNAC. . .)	(POTENTIAL) .
FT	CARBOHYD	1057	1057	N-LINKED (GLCNAC. . .)	(POTENTIAL) .
FT	CARBOHYD	1074	1074	N-LINKED (GLCNAC. . .)	(POTENTIAL) .
FT	CARBOHYD	1081	1081	N-LINKED (GLCNAC. . .)	(POTENTIAL) .
FT	VARIANT	534	534	K -> E (IN ALLOANTIGEN HPA-59;	
FT				dasNP:1801106)	
FT				/FTIG=VAR_003977.	
FT	TURN	170	171		
FT	STRAND	173	180		
FT	TURN	183	184		
FT	HELIX	188	199		
FT	TURN	200	201		
FT	STRAND	204	204		
FT	TURN	206	207		
FT	STRAND	209	216		
FT	TURN	226	224		
FT	TURN	232	228		
FT	HELIX	241	241		
FT	HELIX	252	262		
FT	TURN	263	264		
FT	HELIX	266	268		
FT	TURN	269	269		
FT	STRAND	275	282		
FT	HELIX	289	291		
FT	HELIX	292	301		
FT	TURN	302	303		
FT	STRAND	304	311		
FT	HELIX	313	317		
FT	TURN	318	319		
FT	HELIX	323	330		
FT	TURN	331	332		
FT	HELIX	337	340		
FT	STRAND	341	344		
FT	HELIX	347	353		
FT	HELIX	354	362		
FT	TURN	363	363		
SQ	SEQUENCE	1181 AA;	129295 MW;	7E1B7ED968A94070 CRC64;	
Query Match					17.9%; Score 1054; DB 1; Length 1181;
Best Local Similarity					26.7%; Pred. No. 2.4e-62;
Matches 327; Conservative 216; Mismatches 494; Indels 188; Gaps 43;					
QY	1	FNLDTENAMTPO--ENARGFQSVVQL---	QGRVVVVGAPQEIIVAAQNRGLYQC--DYST 54		
DB	30	YXVGLPEAKIFGSPSEQFGYAVQYQINPKGNWLLVGSFWSGPPENMGVYKCPVDLST 89			
QY	55	GSCETPRLO-----VPVEAVNMSGLSAAATSPQLLACGPTVHOTCSENYVUGLC 107			
DB	90	ATCEKLNLOTSTIPNTEKNTMSGLILTRNMGTFGLTCGLMAQQCGNOYTTGVC 149			
QY	108	FLFGSNLRQOPQPFALRGCPQEDSDIAFLIDGSGIIPHDPFRMKWVSTVMEOLK-- 165			
DB	150	SDISDPF-QLSASFSPATQCPSL-IDVVVVCDESNSIYPWD--AVKNFEXFVQGLDIG 205			
QY	166	KSKTFLSMQYSEFRIHTFFKFNQNPNSRIKXITQLLG-RYTATGLRKYVRELEN 224			
DB	206	PTKTQVGLIQYANNPRVFNLTNTYKTEEMIVATSTQSYGGDLNTTFCGAIQYARKIAYS 265			
QY	225	ITNGAKNAFKILFLTLDGKFDPLGYEDVPELDREGVIRVIGV-----GDAPRSBK 279			

RESULT 15.
ITAG HUMAN
ID ITAG HUMAN STANDARD; PRT; 1167 AA.
AC O75578; Q3UH28;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)

266	DB	AASGRRSATKVVVVVITDGESEH-DGSMLKAVDQCNDHNLRFPGIAGVLGYLNRNALDTKN 324
280	QY	SQELNTVASKPRDHPVQIINNFRALKTIQNLREKIPAEIGTQTGSSSSSEHEMSQEGF 339
325	DB	LIKEIKATIASIPTERYFFNVSDRAALLEKAGTLEQIPFISIEGTVQG-GDNFQEMESQVGF 383
340	QY	SAATISNGP--LLSTVGSYDWAGGVFLYTSKESKSTFINMT--RVSDMDN-DAYLGAYAAAI 394
384	DB	SADYSSQNDIILMLGAVGAFGMSGTIVQKTSHTGLHIFPKQAPDQILQDRHSSYLGYSVAA 443
395	QY	ILNRVQSILVLAGAPRYOHIGLVAMFRONTGMWESNANV-----KTOIGAYFGASLCSV 448
444	DB	ISTGESHTFVAGAPRANTGQIVLYSVN-----ENGNITVIOAHRGDQIGSYFGSVLCSV 498
449	QY	DVDSNGSTDLVLIGAPHYYEQTR--GGQSVCPPLRQGRARQWQCDVLYGEOQOPWGRFG 506
499	DB	DVDKDTITDVLVAGAPMYSDLKXEBGRVYLTIKKGILGQH-----FLEGPEGIENTRFG 555
507	QY	AALTVLGVNGDKLTDVAIGAPGEDNRYGAVLYFHTGSGSISPSHSQRIAGS--KLSPR 564
556	DB	SAIAALSDINMDGFNDVIVGSPLENQNSGAVIYNGHQT-IRTKYSOKILGSDGAFRSH 614
565	QY	LQYFGQSLSGGODLTMDGLVLTGAGQGHVLLRSQPVLRVKAIMEFNPREVARNVFECN 624
615	DB	LQYFGRLDGYDLNGDSITDVSIGAPGVQVQLWSQSIADVAIEASFTEKI--TLVNKN 672
625	QY	DQVKGKEAGEVRVCLHVQKSTRDLREGQIQSVVYDLDL-----SGRPHSRAVFNTEK 680
673	DB	AQ11-----LKLCP-----SAKFRPTKQNNQVAIVYNTILDADGFSRVSRLGFKENN 721
681	QY	NSTRQTOVLGLTQTC--ETLKLQIPNCIEDPVPIVLRINLNFSLVGTPLSAGNLRPVLA 738
722	DB	ERCLQKNVNVQAQSCPEHIIYIOEPS---DVNLSDLRVDISLENPGTS-----PALE 772
739	QY	EDAQRLFTALPPFRKNCNDNTCODDLSITP-----SFMSLDCLVVGGRPREFNVTVVRND 794
773	DB	AYSETAKVFSIPFHKDCGEDGLCISDLVLDVRQIPAAQEQPFIVSNQNKRLTFSTVLKNK 832
795	QY	GEDSVRTQVTFPPDLILSYRKVSTLQNRQSRWELACESAST-EVSGALKSTSCSINH 853
833	DB	RESAYNTGIVVDFSENLPF-----ASFSLPVDGTEVTCQVAASQKSVACDVGY 880
854	QY	PIFPENSEVTENITFDVDSKASLGNKLLKANVTSENNMPRTNKTETFOLELPVKYAVMV 913
881	DB	PALKREQQVTFITNDFNLQ-NLQNASLSPQALSESQENKADNLVNLKIPLLYDAEI- 938
914	QY	VTSHGVSTKYLNFNTASNTSRWQHGYQVSNLGR-----SLPTSLVFLV----- 958
939	DB	---HLTRSTNINFYEISSDGNVPSIVHSFEDVGPKFIFSLKVTGTSVPVSMATVIHIPQ 995
959	QY	-----PVRLNQTVIWRDPQVTP--SENLSSTCHTKER 988
996	DB	YTKRKNPLMYLTGVQTDKAGDISCNADINPLKIGT-----SSSVSPKSENF---HTKB- 1047
989	QY	LPSHSDFLAELKAPVNVCSIAVCORIQCDIPFGIQEEFNATLKNLSFDWTIKTSHNH 1048
1048	DB	-----LNCRATASCNVTCMLKQVHMKGEYFVNVNTRTNWGTFFASTFTQT 1091
1049	QY	LLIVSTAEI-LPNDVSFTLLPGQAFVRSQETKVPPEVNP-----LP-LIVGSSVG 1100
1092	DB	VQLTAASEINTYNEIYVI-----EDNTVTIPLIMKPEDEKAEVPTGVIIGSIIA 1141
1101	QY	GLLLALITAAALYKLGFFKRYKDM 1125
1142	DB	GILLALLAVAILWKLGFPRKRYEM 1166

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Integrin alpha-10 precursor.

GN ITGA10.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Articular chondrocytes;

RX MEDLINE=98352078; PubMed=9885391;

RA Camper L., Hellman U., Lundgren-Aakerlund E.;

RT "Isolation, cloning, and sequence analysis of the integrin subunit

RT alpha10, a beta1-associated collagen binding integrin expressed on

RT chondrocytes";

RL J. Biol. Chem. 273:20383-20389 (1998).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Endothelial cells, and Heart;

RX MEDLINE=20169197; PubMed=10702680;

RA Lehnert K., Ni J., Leung E., Gough S.M., Morris C.M., Liu D.,

RA Wang S.-X., Langley R., Krissansen G.W.;

RT "The integrin alpha10 subunit: expression pattern, partial gene

RT structure, and chromosomal localization.";

RL Cytogenet. Cell Genet. 87:238-244 (1999).

CC -!- FUNCTION: INTEGRIN ALPHA-10/BETA-1 IS A RECEPTOR FOR COLLAGEN.

CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-10

CC ASSOCIATES WITH BETA-1.

CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

CC -!- TISSUE SPECIFICITY: Widely expressed with highest expression in

CC muscle and heart. Found in articular cartilage.

CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS

CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.

CC -!- SIMILARITY: Belongs to the integrin alpha chain family.

CC -!- SIMILARITY: Contains 1 VWFA domain.

CC -!- SIMILARITY: Contains 7 FG-GAP repeats.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AF074015; AAC31952.1; -.

DR EMBL; AF112345; AAF21944.1; -.

DR EMBL; AF172723; AAF61638.1; -.

DR HSSP; P17301; 1A0X.

DR Genew; HGNC:6135; ITGA10.

DR MIM; 604042; -.

DR GO; GO:0008305; C:integrin complex; TAS.

DR GO; GO:0004895; F:cell adhesion receptor activity; TAS.

DR GO; GO:0005518; F:collagen binding; TAS.

DR GO; GO:0007160; P:cell-matrix adhesion; TAS.

DR InterPro; IPR000413; Integrin_alpha.

DR InterPro; IPR002035; VWF_A.

DR Pfam; PF01839; FG-GAP; 3.

DR Pfam; PF00092; vwa; 1.

DR PRINTS; PR01185; VWFADOMAIN.

DR PRINTS; PR00453; VWFADOMAIN.

DR SMART; SM00191; Ict_alpha; 4.

DR SMART; SM00327; VWA_1.

DR PROSITE; PS00242; INTEGRIN_ALPHA; FALSE_NEG.

DR PROSITE; PS0234; VWFA; 1.

KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;

KW Signal; Repeat; Calcium; Magnesium.

FT SIGNAL 1 22

FT CHAIN 23 1167

FT DOMAIN 23 1122

FT TRANSHEM 1123 1145

FT DOMAIN 1146 1167

FT REPEAT 38 97

FT REPEAT 167 350 ? PG-GAP 2.

FT DOMAIN 365 427 ? VWFA.

FT REPEAT 428 482 ? FG-GAP 3.

FT REPEAT 483 545 ? FG-GAP 4.

FT REPEAT 546 605 ? FG-GAP 5.

FT REPEAT 606 660 ? FG-GAP 6.

FT DOMAIN 1134 1140 ? POLY-LEU.

FT CA_BIND 494 502 ? POTENTIAL.

FT CA_BIND 558 566 ? POTENTIAL.

FT CA_BIND 620 628 ? POTENTIAL.

FT DISULFID 76 86 ? BY SIMILARITY.

FT DISULFID 666 675 ? BY SIMILARITY.

FT DISULFID 681 736 ? BY SIMILARITY.

FT DISULFID 789 795 ? BY SIMILARITY.

FT CARBOHYD 98 98 ? N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 234 234 ? N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 336 336 ? N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 364 364 ? N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 733 733 ? N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 763 763 ? N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 839 839 ? N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 921 921 ? N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 1011 1011 ? N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 1018 1018 ? N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 1039 1039 ? N-LINKED (GLCNAC. .) (POTENTIAL).

FT CONFLICT 844 844 ? I -> L (IN REF. 2).

FT CONFLICT 909 909 ? G -> V (IN REF. 2).

FT CONFLICT 926 926 ? E -> D (IN REF. 2).

SEQ SEQUENCE 1167 AA; 127573 MW; AETD3ALC25C1AEAO CRC64;

Query Match 17.9%; Score 1051.5; DB 1; Length 1167;

Best Local Similarity 28.7%; Pred. No. 3,4e-62;

Matches 351; Conservative 199; Mismatches 491; Indels 183; Gaps 45;

QY 1 ENLDTENAMTEQENARG-FGOSVVLQGSRR---VVGAPOSVIAVNOGRSLYQC----- 50

DB 23 ENLDEHPRLPGGPEAEFGYSVLQHVGGGQGWLVGAPWDGPGDRGDVYRCPVGAH 82

QY 51 -----DYSTG-SCEPIRLQVPVEAVNMISLGLSLAFTSPPLACOPTVHQTCE 99

DB 83 NAPCAKHLGIDYQLGNSSHP-----AVNMHLMGMSLLETGDDGGFWACAPLMSRACGS 134

QY 100 NTVYKGLCFEGSNLRQOPQKFPALRGCPQEDSDIAFLIDGSGSIIPHDFFRKEM-V 157

DB 135 SVFSSGICARVDASFPQGSGLAPTAQR-CPYV-MDVVVLVDGSGNSIYP-----MSEV 184

QY 158 STVMEQL-----KSKTFLFSLMOYSEEFRIHFTFKFQNNPNRSLIKPTQLLGR-TH 210

DB 185 QTFLRLVGLFIDPEQIQVGLVOYGSFVHWSLIGDPTKKEVVVRAAKNLSRREGRTK 244

QY 211 TATGLRKVVRELFNITNGARKNAFKILFLTDGKFP-GDPLGYEDVLPDLREGVIRVI 269

DB 245 TAAQIMVACTEGFSQSHGGRPEAAARLLVVDVDSHDEEL--PAALKACBAGVATRYGI 302

QY 270 GV-GDAPRSEKS---ROELNTVASKPPDRHVFOINNFEALKTIONOLREKIFAIEGTOT 324

DB 303 AVLGHYLRQRDPSPLREINTIASDPPERFFNVYDEAALTDIVDALGDRIFGLESHA 362

QY 325 GSSSFHEMSQEGFSAITNSGPELLSTVGSYDWAGGVFLYTSKEKSTPIINTRVDS--- 381

DB 363 ENSSFGLEMSQIGFSTHRLKXGILFGMWGAYDWGGSVLMLEGGHRLPPPRMALEDEFP 422

QY 382 --DMDAYLVYA--NAITLNRVQSVLVGLGAPRYQHIGLVAMER-ONTGWESNANVKGTOI 437

DB 423 ALQNHAAVLGYSVSSMLLGRRLFLSGAPFRHGRKVIAFQKXGAVRVVAQSLOGEQI 482

QY 438 GAYFCASLCSDVIDNSGSTDVLVLGAPHYY--BOTRGQGVSVCPPLPRQORARWQCDVLY 495

DB 483 GSYFGSELCPDTRDGTDTDLVLLVAAPMFLGPNKETGRVYVYL--GQOSLLTLQGTIQ 540

QY 496 GEQGPWGRFGAALTVLGDVNGDKLTQVAIGAPGEENRNGAVLYLPHGTSGLSPSHSOR 555

Db 541 PEPQD-ARPGFANGALPDINODGFADVAVGAPLEDHQCALYLYHGTQ-SGVREHPAQR 598
Qy 556 IAGSKLSPRLOYPGQSLSGQDITMDGLVDLTVGAQHVLRLRSQPLRVKAIMENPRE 615
Db 599 IAAASPHALSYPGRSDGRDLDDGLVDVAVGAQAAILLSRPIVHLTPSLEVTPOA 658
Qy 616 VARNVFECNDQVVKGEAG-EVRVCLHVQKSTRDRLEGQIQSVVTVYDALDSGRPHSR 673
Db 659 ISVVQDCCR--RGQAVCLTALCFQVTSRTPGRWDH---QFYMRPTASLDDEWTAGAR 712
Qy 674 AVNET--KNSTRQTOVTLGTCTCLKQLPNCIEDVPSIVLRLNFSLVGTPLSARG 731
Db 713 AAPDGGQRLSPRLRLSVG-NVTCEQLHPLVD-TSDYLRPVALVTWTFALDNTTKPG-- 768
Qy 732 NLRPLAQAORLFTALPPEKNCNDNICODLSITFSFMSLDC-----LVVGGPR 783
Db 769 ---FVLNBSPTSIOQLNVPSPKDCGPDNECVTLVLQ---VMDIRGSRKAPFVVVGGRR 822
Qy 784 EFNVTVVRNGEDSVYRQVTFPPFLDLVYKVKSTLQNRQSRMSRLACESASSTEVSCA 843
Db 823 KVLVSTTLENKENAYNTSLSIIFSRNL---HLASLTPQR-ESPIKVECAPSA----- 872
Qy 844 LKSTSCSINHPIFPENSEVTENITEDVDSKASLG---NKLL-----LKANVTSENNMPT 895
Db 873 -HARLCSVGHVPFQTKAKVTFLEPEFSCSLUSQVFGKLTASSDSLERNGTLQENTACT 931
Qy 896 NKTEFOLELPLVAVYVMTVSHGVSTKYLMTASENTSRYMCHQY----- 940
Db 932 -----SAYIQYEPH-----LLFSSESTLHRYEVHPYCTLPVGPGRPKTTL 972
Qy 941 QVSNLG---QRSLPISLVPLV-----VRLNQTVIWDROPQVTFSENLSSTCHTKER 988
Db 973 RVQNLGCVYVSGLIISA--LLPAVARGNYPUSLSQVI-----TNASCIVQNLTE 1021
Qy 989 LPSHSDFLAELRKAPVWVNCIAVCQRIQCDIPFGIQBEFNATLKGNLSDWDYIKTSHNH 1048
Db 1022 PPGPPVHPELOHTNRLNGSNTQCVVRCHLGOLANGTEVSVCLLRLVNEHERRAKFS 1081
Qy 1049 LLIVSTABILLFNDVSFTLLPGQGAFTVRSOTETKVBPPEVFNPLPLIVGSSVQGLLALLI 1108
Db 1082 LTVVSTFELGTERGSLVQLTEASRWSESLLEV-VQTRPILISLWILIGSVLQGLLALL 1140
Qy 1109 TAALYKLGPF-----KROYK 1123
Db 1141 VFCLWKLGPFAHKKIPEERKREK 1164

Search completed: June 7, 2004, 17:12:47
Job time : 14.9719 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 16:56:30 ; Search time 39.4889 Seconds

(without alignments)
9084.693 Million cell updates/sec

Title: US-09-902-481B-4

Perfect score: 5884

Sequence: 1 FNLDTENAMTFQENARGFQ.....PKRQYKDMMSGGPPGAEPPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 25:*

- 1: sp.archaea.*
- 2: sp.bacteria.*
- 3: sp.fungi.*
- 4: sp.human.*
- 5: sp.invertebrate.*
- 6: sp.mammal.*
- 7: sp.mhc.*
- 8: sp.organelle.*
- 9: sp.phage.*
- 10: sp.plant.*
- 11: sp.rodent.*
- 12: sp.virus.*
- 13: sp.vertebrate.*
- 14: sp.unclassified.*
- 15: sp.virus.*
- 16: sp.bacteriap.*
- 17: sp.archaea.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	4371	74.3	1151	11 Q9J130	Q9J130 rattus norv
2	3891.5	66.1	1036	11 Q8CA73	Q8CA73 mus musculus
3	3759.5	64.6	920	6 Q28984	Q28984 sus scrofa
4	3470	59.0	1169	4 Q81VAG	Q81VAG homo sapien
5	3310.5	56.3	1169	11 Q9QXH4	Q9QXH4 mus musculus
6	3236.5	55.0	1161	11 Q9QYB7	Q9QYB7 rattus norv
7	1523.5	25.9	1161	11 Q9WTV4	Q9WTV4 mus musculus
8	1513	25.7	1160	11 Q9R200	Q9R200 mus musculus
9	1395	23.7	1196	13 Q98TF1	Q98TF1 cyprinus ca
10	1350.5	23.0	1086	4 Q96HB1	Q96HB1 homo sapien
11	1344.5	22.9	1187	13 Q98TF0	Q98TF0 cyprinus ca
12	1269	21.6	927	6 Q8HZV0	Q8HZV0 bos taurus
13	1154.5	19.6	1167	11 Q88340	Q88340 rattus norv
14	1109	18.8	1167	11 Q88341	Q88341 rattus norv
15	1043.5	17.7	1188	11 Q7QC3	Q7QC3 mus musculus
16	1041	17.7	1171	13 Q42094	Q42094 gallus gall

17	1032	17.5	1038	11 Q8BS01	Q8BS01 mus musculus
18	1005	17.1	895	11 Q9WUF8	Q9WUF8 mus sp. itg
19	998.5	17.0	1160	6 Q8MK24	Q8MK24 felis silve
20	865	14.7	348	4 Q8TES5	Q8TES5 homo sapien
21	860	14.6	1332	5 Q8BPQ8	Q8BPQ8 halocynthia
22	796	13.5	205	11 Q63001	Q63001 rattus norv
23	753.5	12.8	780	13 Q06271	Q06271 xenopus lae
24	738	12.5	823	4 Q8WY18	Q8WY18 homo sapien
25	686.5	11.7	823	11 Q8CE84	Q8CE84 mus musculus
26	669	11.4	1032	11 Q61989	Q61989 mus musculus
27	644.5	11.0	1033	6 Q9BGU3	Q9BGU3 bos taurus
28	643	10.9	1036	11 Q91VD5	Q91VD5 mus musculus
29	623.5	10.7	1474	5 Q86G87	Q86G87 pseudoplusi
30	628.5	10.7	257	11 Q8C270	Q8C270 mus musculus
31	619.5	10.5	1041	5 Q9UB90	Q9UB90 lytechinus
32	614.5	10.4	1041	5 Q76378	Q76378 lytechinus
33	580	9.9	1034	13 Q98VT7	Q98VT7 gallus gall
34	579.5	9.8	1054	5 Q9UGS1	Q9UGS1 strongyloce
35	555.5	9.4	1053	11 Q80YP5	Q80YP5 mus musculus
36	550	9.3	1033	13 Q42598	Q42598 xenopus lae
37	546	9.3	1036	6 Q7YRP8	Q7YRP8 equus cabal
38	534	9.1	1016	13 Q91779	Q91779 xenopus lae
39	530	9.0	974	11 Q924W2	Q924W2 rattus norv
40	529	9.0	1073	11 Q8CC06	Q8CC06 mus musculus
41	527.5	9.0	1119	5 Q86G88	Q86G88 pseudoplusi
42	526	8.9	1047	6 Q9WZD6	Q9WZD6 bos taurus
43	525.5	8.9	1007	6 Q9GK48	Q9GK48 bos taurus
44	522.5	8.9	1132	11 Q80Z18	Q80Z18 mus musculus
45	512.5	8.7	1034	6 Q9TUN4	Q9TUN4 oryctolagus

ALIGNMENTS

RESULT 1

Q9J130 PRELIMINARY; PRT; 1151 AA.
AC Q9J130;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Integrin beta 2 alpha subunit.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Fathallah D.M. Sr., Zerrin K. Jr.;
RT "Cloning of the rat CD11b cDNA sequence."
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF268593; AAF81280.1; -
DR HSP; P11215; 1BHQ.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PRO1185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWFA; 1.
SQ SEQUENCE 1151 AA; 126943 MW; 8F785695D4074CA5 CRC64;

Query Match 74.3%; Score 4371; DB 11; Length 1151;
Best Local Similarity 72.6%; Pred. No. 1.9e-313;
Matches 826; Conservative 150; Mismatches 159; Indels 2; Gaps 2;

```
QY 1 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVGAPQBIIVANQBSGLYQCDYSTGSCBPI 60
Db 17 FNLDTEHPMTFOENARGFGQSVVQLQGSRRVVGAPQBIIVANQBSGLYQCDYSTGSCBPI 76
QY 61 RLQVPEAVNNSLGLSLAAATTPPQALLACGPTVHQCKENTYVNGCLYFGSNLLRPQ 120
Db 77 PLQVPEAVNNSLGLSLAAATTPPQALLACGPTVHQCKENTYVNGCLYFGSNLLRPQ 136
QY 121 FPEALRGCPQSDSDIAFLIDGSGSIIPHDPRFMKEWSTVMEQKSKTFLSLMOYSEF 180
Db 137 FPEALRGCPQSDSDIAFLIDGSGSIIPHDPRFMKEWSTVMEQKSKTFLSLMOYSEF 196
QY 181 RIHFTFKFONNPNRSLIKPITOLLGRTHATGLRAXVRELEWITNGARKNAFKILFL 240
Db 197 RIHFTFKFONNPNRSLIKPITOLLGRTHATGLRAXVRELEWITNGARKNAFKILFL 256
QY 241 TDGSKFGDPLGYEDVPELDRGVRVIVGVGDAFRSEKSRQELNVTASKPPDRHVQIN 300
Db 257 TDGSKFGDPLGYEDVPELDRGVRVIVGVGDAFRSEKSRQELNVTASKPPDRHVQIN 316
QY 301 NFALKTIONLRKIPAIETGQTGSSSPHEMSQEGFSAATISNGPILLSVGSYDWDAG 360
Db 317 NFALKTIONLRKIPAIETGQTGSSSPHEMSQEGFSAATISNGPILLSVGSYDWDAG 376
QY 361 GVFLYTSKZKSTFINNTRVDSMDNDAYLGVAALILNLRVOSLVLAGPRYOHIGLVAMER 420
Db 377 GAFLYPSKDKASFINNTRVDSMDNDAYLGVAALILNLRVOSLVLAGPRYOHIGLVAMER 436
QY 421 QNTGMWESNANVKTQIYGAFGLSCVDVDSNGSDTLVIGAPHYYBOTRGQGVSCPL 480
Db 437 QNTGMWESNANVKTQIYGAFGLSCVDVDSNGSDTLVIGAPHYYBOTRGQGVSCPL 496
QY 481 PRGORARQWQDAVLYGSGQFMRFGAALTVLGVDVNGDKLTVIAGPGEEDNRGAVYLF 540
Db 497 PRGORARQWQDAVLYGSGQFMRFGAALTVLGVDVNGDKLTVIAGPGEEDNRGAVYLF 555
QY 541 HGTSGSGISPSHQRAGSKSLRLOVFGQSLGGQDLTMDGLVDTLVGAQGHVLLRQ 600
Db 556 HGTSGSGISPSHQRAGSKSLRLOVFGQSLGGQDLTMDGLVDTLVGAQGHVLLRQ 615
QY 601 PVLRVKAIMEFNEPREVARNVFCNDQVVKKEAGEVRVCLVHVKSTRDLREGQIOSVT 660
Db 616 PVLRVKAIMEFNEPREVARNVFCNDQVVKKEAGEVRVCLVHVKSTRDLREGQIOSVT 675
QY 661 YDLALDSGRPHSRAVNFKSTRQTVGLTQCTETKLQPCNCEIDPVSIVRLNPF 720
Db 676 YDLALDSGRPHSRAVNFKSTRQTVGLTQCTETKLQPCNCEIDPVSIVRLNPF 735
QY 721 SLVGTPLSAPGNLRPVLAEADQRLTALPPEKNCQNDNICDDLSITPFSMSLCLVWG 780
Db 736 SLVGTPLSAPGNLRPVLAEADQRLTALPPEKNCQNDNICDDLSITPFSMSLCLVWG 795
QY 781 GPREFNVTVVRNDGDSYQTVFTFFPLDLSVRKVSTLQNSQSRWLAACESASTEV 840
Db 796 GPREFNVTVVRNDGDSYQTVFTFFPLDLSVRKVSTLQNSQSRWLAACESASTEV 854
QY 841 SGALKSTSCSINHPIIPENSEVFNITFDVDSKASIGNKLLKANTVSENMPRTNKTFF 900
Db 855 SGALKSTSCSINHPIIPENSEVFNITFDVDSKASIGNKLLKANTVSENMPRTNKTFF 914
QY 901 QLELPVKYAVMYVTSHGVSTKYLNTFASNTSRVMOHQYQVSNLQORSLPISLVELVPV 960
Db 915 QLELPVKYAVMYVTSHGVSTKYLNTFASNTSRVMOHQYQVSNLQORSLPISLVELVPV 974
QY 961 RLNQTVMDRPOVTFENISSTCHTERLPSHSDLAELRKAPVUNCSTAVQORLQCDLP 1020
Db 975 RLNQTVMDRPOVTFENISSTCHTERLPSHSDLAELRKAPVUNCSTAVQORLQCDLP 1034
QY 1021 FFGIOEFNATLKNLSFDWYIKTSNNHLLIVSTABIFLNDVSYFTLLPQCGAFVRSQTE 1080
Db 1035 FFGIOEFNATLKNLSFDWYIKTSNNHLLIVSTABIFLNDVSYFTLLPQCGAFVRSQTE 1094
QY 1081 KVEPPEVPNPLPLIVGVSSVGLLILLALITAAALKLGFRRQYKDMMSGEGPFGAEPQ 1137
```

```
Db 1095 KVEPPEVPNPLPLIVGVSSVGLLILLALITAGLYKLGFFRRQYKDMMSGEGPFGAEPQ 1151
RESULT 2
Q8CA73 PRELIMINARY; PRT; 1036 AA.
AC Q8CA73;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Integrin alpha M.
GN ITGAM OR P730045J24RIK.
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spinal cord;
RX MEDLINE=22354683; PubMed=12466851;
RA THE PANTOM Consortium,
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT -Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.;
RL Nature 420:563-573 (2002).
DR EMBL; AK039444; BAC30350.1; -.
DR PIR; PT0572; PT0572.
DR PIR; PT0633; PT0633.
DR PIR; PT0697; PT0697.
DR MGD; MGI:96607; Itgam.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 1.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS02334; VWFA; 1.
SQ SEQUENCE 1036 AA; 115126 MW; 4F4F39BF1E188B77 CRC64;
Query Match 66.1%; Score 3891.5; DB 11; Length 1036;
Best Local Similarity 66.0%; Pred. No. 4.3e-278;
Matches 751; Conservative 125; Mismatches 143; Indels 119; Gaps 2;
```

```
QY 1 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVGAPQBIIVANQBSGLYQCDYSTGSCBPI 60
Db 17 FNLDTEHPMTFOENARGFGQSVVQLQGSRRVVGAPQBIIVANQBSGLYQCDYSTGSCBPI 76
```

```
QY 61 RLQVPEAVNNSLGLSLAAATTPPQALLACGPTVHQCKENTYVNGCLYFGSNLLRPQ 120
Db 77 PLQVPEAVNNSLGLSLAAATTPPQALLACGPTVHQCKENTYVNGCLYFGSNLLRPQ 136
```

```
QY 121 FPEALRGCPQSDSDIAFLIDGSGSIIPHDPRFMKEWSTVMEQKSKTFLSLMOYSEF 180
Db 137 FPEALRGCPQSDSDIAFLIDGSGSIIPHDPRFMKEWSTVMEQKSKTFLSLMOYSEF 196
```

```
QY 181 RIHFTFKFONNPNRSLIKPITOLLGRTHATGLRAXVRELEWITNGARKNAFKILFL 240
Db 197 RIHFTFKFONNPNRSLIKPITOLLGRTHATGLRAXVRELEWITNGARKNAFKILFL 256
```

```
QY 241 TDGSKFGDPLGYEDVPELDRGVRVIVGVGDAFRSEKSRQELNVTASKPPDRHVQIN 300
Db 257 TDGSKFGDPLGYEDVPELDRGVRVIVGVGDAFRSEKSRQELNVTASKPPDRHVQIN 316
```

```
QY 301 NFALKTIONLRKIPAIETGQTGSSSPHEMSQEGFSAATISNGPILLSVGSYDWDAG 360
Db 317 NFALKTIONLRKIPAIETGQTGSSSPHEMSQEGFSAATISNGPILLSVGSYDWDAG 376
```

```
QY 361 GVFLYTSKEKSTINMTRVDSNDNDAYLVAAIILRNVRQSVLVGAPRYQHIGLVAMFR 420
DB 377 GAFLYTSKOKVITINTRVDSNDNDAYLVASAVILENRVQSVLVGAPRYQHIGLVAMFR 436
QY 421 QNTCMFESNANVGTQIGAFGASCLSDVDNMGSDLVGAPRYQHIGLVAMFR 480
DB 437 ENFGTWPHTSIK-----450
QY 481 PRQORARWQDAVLYGEGQPGWGRFGAALTVLGDNVNGDKLTDVAIGAPGEEDNRGAVYLF 540
DB 451 -----450
QY 541 HGTSGSGISPHSORIAGSKLSPRIQYFGOSLSGGDLTMDGLVLTGVAQGHVLLRSQ 600
DB 451 -----SORIAGHFSFQYFGOSLSGGDLTMDGLVLTGVAQGHVLLRSQ 498
QY 601 PVLVRKALMEFNPREVARNVFCNDQVVGKEAGEVRVCLHWQKSTRDLREGIQSVVT 660
DB 499 PVLRLBATMEFSPKRVARSVFACQEQVLKNDAGEVRVCLRWKNTKRLREGIQSVVT 558
QY 661 YDLALDSGRPHSAVNETKSTRROTQVGLTCTCETLKLQLPNCIEPVPVILRLNP 720
DB 559 YDLALDSGRPHSAVNETKSTRROTQVGLTCTCETLKLQLPNCIEPVPVILRLNP 618
QY 721 SLVGTPLSAFGLNRPVLAEDAQRLEFALPFFERNKCGNDNICODDLSTIFFSMGLDCLVVG 780
DB 619 TLVGEPLRSFGLNRPVLAEDAQRLEFALPFFERNKCGNDNICODDLSTIFFSMGLDCLVVG 678
QY 781 GRPEENVTVVRNDGDSYRTQVTFPFLDLSVRKVSTLQNRQSRWBL-ACESASSTE 839
DB 679 GPQDFNMSVTLRNDGDSYRTQVTFPFLDLSVRKVSTLQNRQSRWBL-ACESASSTE 738
QY 840 VSCALSTSCSINHPIFFENSEVFNITFDVDSKASLGKLLKANKVTSNNPRTNKTE 899
DB 739 GHICALSTTWINHPIFFANSEVFNITFDVDSKASLGKLLKANKVTSNNPRTNKTE 798
QY 900 FQLELPVKYAVVWTSVSHGVTYKILNFTASENTRVQVQVQVQVQVQVQVQVQVQV 959
DB 799 FQLELPVKYAVVWTSVSHGVTYKILNFTASENTRVQVQVQVQVQVQVQVQVQVQV 858
QY 960 VRLNQTWIDRPOVTFSENLSSTCHTKERLPHSDFLAELKAPVNVCSIAVCRIOQCDI 1019
DB 859 VQIINNVTWIDRPOVTFSENLSSTCHTKERLPHSDFLAELKAPVNVCSIAVCRIOQCDI 918
QY 1020 PFGIOEENFATLKNLSFDWIKTSHNHLIIVSTABILLFNDSVFTLLPQOGAFVRSQTE 1079
DB 919 PSNTQEIENFATLKNLSFDWIKTSHNHLIIVSTABILLFNDSVFTLLPQOGAFVRSQTE 978
QY 1080 TKVEPEVNPPLVIGSSVGLLALITAILYKLGFEKROYKDMWSEGGPPGAPQ 1137
DB 979 TKVEPEVNPPLVIGSSVGLLALITAILYKLGFEKROYKDMWSEGGPPGAPQ 1036

RESULT 3
Q28984
ID Q28984 PRELIMINARY; PRT; 920 AA.
AC Q28984;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-FEB-1997 (TREMblrel. 02, Last sequence update)
DT 01-JUN-2003 (TREMblrel. 24, Last annotation update)
DE CD11b (Fragment).
GN CD11B.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RA SEQUENCE FROM N.A.
RA Lee J.-K., Schook L.B., Rutherford M.S.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
DR ENBL; U00072; AAB16869.1; -.
DR HSSP; P11215; 1BHQ.
```

```
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS50234; VWEA; 1.
FT NON_TER 1
FT SEQUENCE 920 AA; 102440 MW; E96CC518350DD5AC CRC64;
QY 118 POKFPEALRGCPQEDSDIAFLIDGSGSIIIPDPRMKEMVSTVMEQLKSKTLPISLMQYS 177
DB 1 POKFPEALRGCPQEDSDIAFLIDGSGSIIIPDPRMKEMVSTVMEQLKSKTLPISLMQYS 60
QY 178 EEFRIHPTKFEQNNPNRSLIKPITQLLGRTHATGLRVVRELFNITNGARKNAFKIL 237
DB 61 EDFYTHPTFENDPKRNPSEKLLVPEIRQLLGRTHATGIRKVVRELFHSGSGARENAKIL 120
QY 238 FLITDGEKFGDPLGVEDVIPELDRGVIRYVIGVDGAFRSEKSRQELNATVASKPRDHVF 297
DB 121 VVITDGEKFGDPLGVEDVIPEADRGVIRYVIGVDGAFRSEKSRQELNATVASKPRDHVF 180
QY 298 QINNPEALKTKQNLREKIFAIETGQTSSSSFEHMSQEGFSAITNGPILSTVGSYD 357
DB 181 QVNNPEAVKTIQNLQKTEFALGQTGTSSTFCEMGEQFSAITNGPILGAVGSD 240
QY 358 WAGGVFLTSYKSTKSTFNNTRVDSMDNDAYLVAAIILRNVRQSVLVGAPRYQHIGLV 417
DB 241 WAGGAFLEMPKDRVIFINTTRVDSMDNDAYLVAAIILRNVRQSVLVGAPRYQHIGLV 300
QY 418 MFRONTGMWSENANVKGTOIGAYFGASLCSDVDNMGSDLVGAPRYQHIGLV 477
DB 301 MFKQNSGAWENKADIKSGIISYFGASLCSDVDNMGSDLVGAPRYQHIGLV 360
QY 478 CPLRQGRARWQDAVLYGEGQPGWGRFGAALTVLGDNVNGDKLTDVAIGAPGEEDNRGAV 537
DB 361 CPLRQGRARWQDAVLYGEGQPGWGRFGAALTVLGDNVNGDKLTDVAIGAPGEEDNRGAV 419
QY 538 YLEHGTSGSGISPHSORIAGSKLSPRIQYFGOSLSGGDLTMDGLVLTGVAQGHVLL 597
DB 420 YLEHGTSGSGISPHSORIAGSKLSPRIQYFGOSLSGGDLTMDGLVLTGVAQGHVLL 479
QY 598 RSQPVLRVYKATMEFNPREVARNVFCNDQVVGKEAGEVRVCLHWQKSTRDLREGIQS 657
DB 480 RSQPVLRVYKATMEFNPREVARNVFCNDQVVGKEAGEVRVCLHWQKSTRDLREGIQS 539
QY 658 VVTDLALDSGRPHSAVNETKSTRROTQVGLTCTCETLKLQLPNCIEPVPVILRLNP 717
DB 540 VVTDLALDSGRPHSAVNETKSTRROTQVGLTCTCETLKLQLPNCIEPVPVILRLNP 599
QY 718 LNFSLVGTPLSAFGLNRPVLAEDAQRLEFALPFFERNKCGNDNICODDLSTIFFSMGLDCL 777
DB 600 LNFSLVGTPLSAFGLNRPVLAEDAQRLEFALPFFERNKCGNDNICODDLSTIFFSMGLDCL 659
QY 778 VVGSPRENVTVVRNDGDSYRTQVTFPFLDLSVRKVSTLQNRQSRWBL-ACESASSTE 837
DB 660 VVGSPRENVTVVRNDGDSYRTQVTFPFLDLSVRKVSTLQNRQSRWBL-ACESASSTE 719
QY 838 TEVSGALKSTSCSINHPIFFENSEVFNITFDVDSKASLGKLLKANKVTSNNPRTNKTE 897
DB 720 TEVSGALKSTSCSINHPIFFENSEVFNITFDVDSKASLGKLLKANKVTSNNPRTNKTE 779
QY 898 TFOLELPVKYAVVWTSVSHGVTYKILNFTASENTRVQVQVQVQVQVQVQVQVQVQV 957
```

```
Query Match 64.6%; Score 3799.5; DB 6; Length 920;
Best Local Similarity 78.8%; Pred. No. 2.2e-271;
Matches 726; Conservative 84; Mismatches 110; Indels 1; Gaps 1;
```

```

780 TEFQLEPVKAVYVWVTSLEVSATSEKXTRHVEHIOYQFNNLQCKLPISVWF 839
958 VPEVLNQTIVNDPQVTFSENLSTCTKELPSHSDFLAELKAPVNVCSIAVCORIC 1017
840 VPVLRNVTVWDQVQVTFSQLSRCSCTEELGPRHSDFLKQLKTPVINCIAVCQKIC 899
1018 DIPFGI:QSEFNATLKNLSF 1038
900 DIPSGI:QSEELKVTLKNLSF 920

RESULT 4
Q81VA6 PRELIMINARY; PRT; 1169 AA.
AC Q81VA6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Integrin, alpha X (antigen CD11C (p150), alpha polypeptide).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RA Strausberg R.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC038237.1; -.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWFA; 1.
SQ SEQUENCE 1169 AA; 128521 NW; A17B484FEC79EB6 CRC64;

```

```

Query Match 59.08; Score 3470; DB 4; Length 1169;
Best Local Similarity 61.08; Pred. No. 7.2e-247;
Matches 689; Conservative 142; Mismatches 292; Indels 6; Gaps 4;

```

```

Qy 1 FNLDENAMTFQENARGFGSVVQLQGSRRVVVGAPOBIVAAVNRGSLYQCDYSTGSCPEI 60
Db 20 FNLDTEELTAFRVDSAGFGSDSVVQYANSVVVGAPOKI TAANQTGGLYQCYSTGACEPI 79
Qy 61 RLQVPEAVNMSLGLSLAATTSPPOLLAGPVTHTQTCSENTYVKGCLPFGSNLRQOPQK 120
Db 80 GLQVPEAVNMSLGLSLAATTSPPOLLAGPVTHTQTCSENTYVKGCLPFGSNLRQOPQK 137
Qy 121 PPEALRGCPQSDSIAFLIDGSGIIPHDPRRMKEWSTVWBLKKSCTLFLSLMOYSEBF 180
Db 138 LPVSRQCPQEQDIFVLIDGSGISSENATMNFVRAVISQFQRPSTQFSLMQFSNKP 197
Qy 181 RIHTPFKEQNNPNRSLIKPITQLGRTHATGLRKVKVRELNITNGARKNAFKILFL 240
Db 198 QHTPTFEFRSSNPLSILASVHQLQGYTATAIQNVHRLFEASGARDAKILIVI 257
Qy 241 TDGKFGDPLGYENVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVPQIN 300
Db 258 TDGKFGDPLGYENVIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVPQIN 317
Qy 301 NFEALTIQNLREKIFAIEGTQTGSSSPHEHMSQGFSAAITNSGPLLSTVGSYDWAG 360
Db 318 DFDALKIQNLQKEKIFAIEGTQTGSSSPHEHMSQGFSAAITNSGPLLSTVGSYDWAG 377

```

```

Qy 361 GVPLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSVILCAPRYQHIGLVAMER 420
Db 378 GAPLYPPNNSPTFINNSQENVDWDSYLGSTELAKWGVQSVILCAPRYQHTGKAVIPT 437
Qy 421 QNTGEMESNANVKGTOIGAYFGASLCSDVDNSGSDTLVLIGAPHYETQTRGQNSVCPL 480
Db 438 QVSRQRMKAEVGTGIGSYFGASLCSDVDNSGSDTLVLIGAPHYETQTRGQNSVCPL 497
Qy 481 PRGORARMOCDVLYGEOQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db 498 PRGWR-RWCDVLYGEOQPMGRFGAALTVLGDVNGDKLTDVIGAPGEENRGAVYLF 556
Qy 541 HGTSGGISPSHSQRIAGSKLSPRLOYFQOSI SGGODLTMDGLDVLDTVGAQHVLIRSQ 600
Db 557 HGVLGPSISPSHSQRIAGSKLSPRLOYFQOALSGGODLTQDGLVDLAVGARGQVLLRTR 616
Qy 601 PVLRYKALMEFNPVARNVFCNDQYKVGKEAGEVVCVCHVOKSTRDLREGQIQSVT 660
Db 617 PVLWGVSNQFIPAEIPRPAFCRCQVVSQETLVQSNICLYIDKRSKNLLSGRDLQSSVT 676
Qy 661 YDLALDSGRPHSRVFNETNSRRQTQVIGLTQTCETLKLQPLNCIEBPVSPVILRNF 720
Db 677 LDALDPGRLSPRAVTFQETKRSLSRVRVLGLKHCENFNILLPSVCDVSVPITLRLNF 736
Qy 721 SLVGTPLSAPGNLRPVLAEDAQLTALPPEKNCNDMI CODDLSITESPMSLDCLVVG 780
Db 737 TLVGPPLAFLRNLRPMLAADAQRYFTASLFFPKNCGDHICQDNLGISFSPGLKSLVVG 796
Qy 781 GPRFNVTVVRNDGEDSYRTQVTFPFLDLSTYRKVSTLQNSORSQSWELACESASSTEV 840
Db 797 SNLELNAEVMWWDGEDSYGTTVTFSHPAGLSYRVVAEGQKQLRSLHLTCDSPVGV-- 854
Qy 841 SGALKSTSCSINHPIPPENSEVTNITFDVDSKASLGKLLKANKVTSNNRPTKTFP 900
Db 855 SQGTWSTSCRINHLIFRGAQITFLATPDVSPKAVLGDRLLLTANVSSNNTPTSKTFP 914
Qy 901 QLELPVKYAVYVWVTSLEVSATSEKXTRHVEHIOYQFNNLQCKLPISVWF 959
Db 915 QLELPVKYAVYVWVTSLEVSATSEKXTRHVEHIOYQFNNLQCKLPISVWF 974
Qy 960 VRLNQVTVNDPQVTFSENLSTCTKELPSHSDFLAELKAPVNVCSIAVCORIC 1019
Db 975 VELNQEAVMQVDESHQPNFSLRCSCTEELGPRHSDFLKQLKTPVINCIAVCQKIC 1034
Qy 1020 PFEQGEENATLKNLSFDTWIKTSHNLLIVSTAEILFNDSTVFTLLPQCAFVSQTE 1079
Db 1035 PFSVQGEELDTLKNLSFQWVVRQILQKVSVVSVVAEITFDSVTSQLEPQCAFVMAQT 1094
Qy 1080 TKVEPFEVNPFLTVGSSVGLLILALITAAALYKLGPFKQYKDMSE 1128
Db 1095 TVLEKYKVHNPFLTVGSSVGLLILALITAAALYKLGPFKQYKDMSE 1143

RESULT 5
Q9QXH4 PRELIMINARY; PRT; 1169 AA.
AC Q9QXH4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Leukocyte adhesion glycoprotein p150.95 alpha integrin subunit.
GN ITGAX.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Huang X., Gorski K., Tong C., Rattis F.-M., Tseng S.-Y., Pardoll D.,
RA Tsuchiya H.;
RT "Isolation of Genes Selectively Expressed by Dendritic Cells.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF211864; AAF23492.1; -.

```

```
DR HSSP; P11215; 1BHQ.
DR MGD; MGI:96609; Itgax.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR GO; GO:0007229; P:integrin-mediated signaling pathway; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; EG-GAP; 3.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS50234; VWA; 1.
KW Integrin.
SQ
Query Match      56.3%; Score 3310.5; DB 11; Length 1169;
Best Local Similarity 56.6%; Pred. No. 4.4e-235;
Matches 645; Conservative 173; Mismatches 303; Indels 19; Gaps 7;

QY 1 FNLDTENAMTFOENARGFQSVVQLQGSRRVVGAPQEIIVAAHQSGLSYQCYSTGSCPEI 60
Db      :      :      :      :      :      :      :      :      :      :
20 FNLDAEKLTHFMDGAEFQSHSLQVYDSSWVVGAPKEIKATNQIGGLYKGYHTGNCPEI 79
QY 61 RLQVPEAVNMSIGLSLAATTSPOLLACGPTVHOTCSNTVYVKGCLFLGCSNLRQOQK 120
Db      :      :      :      :      :      :      :      :      :      :
80 SLQVPEAVNISLGSLSAATNPMLLACGPTVHTCRENIYLTGLCLFLSSFFQS-QN 138
QY 121 FPEALRGCPQEDSDIAPLIDGSGSIIPHDFRMKEMVSTVMBQLKSKTLFLSMQYSEEF 180
Db      :      :      :      :      :      :      :      :      :      :
139 FPTAQCECPKQDQIVFLIDGSGSIISSTDFEKLDFVKAQVMSQLQRPSTFSLMQFSDYF 198
QY 181 RHFTFKFQNNPNSRIKPIITOLLGRTHATGLRKVVRELFNITNGARKNAFKILFL 240
Db      :      :      :      :      :      :      :      :      :      :
199 RVHFTFNNPISSTPLSLGSRQLRGYTTTASAKHVTITLFTTQSGARQDQATKVLVI 258
QY 241 TDGEKFGDPLGYEDVIPELDREGRVYVIGVGDGAFRSKSRQELMTVASKPRDRVFGQIN 300
Db      :      :      :      :      :      :      :      :      :      :
259 TDGRKQGNLSYDSVTPMAEASIRYALGVGKAFYNEHSKQELAKIASMPSEHVFSVE 318
QY 301 NFEALKTONQUREKIFAIEGHTQSSSSFEHMSQEGFSAAITNGPLSTVSGYDWAQ 360
Db      :      :      :      :      :      :      :      :      :      :
319 NFDALKDIEHQLEKIPAEIGETETPSSSTFELEMSQEGFSAVFTDPDGVPLGAVGFSWSG 378
QY 361 GVFLYTSKESKTFINMTRVDSMDAYLGYAAAIILNRVQSLVLGAPRYOHIGLVAMFR 420
Db      :      :      :      :      :      :      :      :      :      :
379 GAFLYPSNMRPTFINMSQENEDRDAYLGYSTALAFWKGVHSLILGAPRHQHTGKVIFT 438
QY 421 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEYQTRGGQVSVCP 480
Db      :      :      :      :      :      :      :      :      :      :
439 QESREHWRPKSEVRGTQIGSYFGASLCSVDMDRSGTDLVLIGVPHYEYHTRGGQVSVCPM 498
QY 481 PRGQARWOCDAVLGEGCOQPGRGGAALTVLGDVNGDKLTDVAIGAGEEDNRCAYVLF 540
Db      :      :      :      :      :      :      :      :      :      :
499 P-GVGRHWCQHTLGEQHPGRFGAALTUVDGVNGDSLADVALGAPGEENRGAVYIF 557
QY 541 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGGQDLTMDGLVDLTVGAQCHVILLASQ 600
Db      :      :      :      :      :      :      :      :      :      :
558 HCAQRQDIAPSPQISASQIPSRIOYFGQSLSGGQDLTRDGLVDLAVGSKGRVLLNTR 617
QY 601 PVLRYKALMEPNPREVARNVFECNDVYKGEAGEVRVCLHVOKSTRDRLEGQIOSVVT 660
Db      :      :      :      :      :      :      :      :      :      :
618 PLRVSVPVTHFPASISRSVFECQVAPQEQFLSDATVCLHSHSPKQL--GDLRSVTI 675
QY 661 YDLALDSGRPHSRVAFNFKTSRQTQVGLHTQTCETLKLQLPNCIEDPVSPIVLRNLF 720
Db      :      :      :      :      :      :      :      :      :      :
676 FDLALDHGRSLSTRAIFKTKTRALTRVTLGLNKHCEVSKLLPACVEDSVTPITLRNLF 735
QY 721 SLVGTPLSFAFNLRPVLAEDAQRLLTALFPPEFKNGCNDNICQDDLSTTFPSMSLDCLVVG 780
```

```
Db      :      :      :      :      :      :      :      :      :      :
736 SLVGVPISSEQLNQMLAVDDQTYFTASLPFEKNGADHICQDDLSVVFPGPDLTKTLVVG 795
QY 781 GPRENVTVTVVRNDCGDSYRTQVTFFPPLDLSSYKXVSTLQ-----NQRQSGWR 829
Db      :      :      :      :      :      :      :      :      :      :
796 SDLELVNDVTVSNDGDSYGTIVTLFYPVGLSPRRVABGQVFLRKCKEDQQRRCQHSILH 855
QY 830 LACASASSTEVSGALKKSTSCSINHIPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSE 889
Db      :      :      :      :      :      :      :      :      :      :
856 LMCD--STPDRSQGLWSTSCSSRHVIFRGGSQMTFLVTFDVSPPKABLDRLLLRARVGE 913
QY 890 NMPRTNKTETQLELPVKYAVYVMVTSRGVSTKVLNFTASB-NTSRVMQHQVQVSNLGR 948
Db      :      :      :      :      :      :      :      :      :      :
914 NNVPCTPKTTFQLELPVKYAVYTWISSHDQTKVLNFTSEKETSVEHRFQVNNLQOR 973
QY 949 SLPISLVLPVRLNQTVINDRPOVTFSENLSSTCHTKERLPSSHDFIAELRKAPVNCVS 1008
Db      :      :      :      :      :      :      :      :      :      :
974 DVPVSINFVPIELKGEAVW-TVMVSHPNQLTCYRNRLKPTQFDLLTTHMOKSVLDCS 1032
QY 1009 IAVCORIOCDIPFFGIOEFNATLKGNLSPDWYIKTSHNLLIYSTAILFNDSVFTLLP 1068
Db      :      :      :      :      :      :      :      :      :      :
1033 IADCLHRCDIPSGILDLYFLKGNLSFGWISQTLQKKVLLISEAITEFTNTSVYSQLP 1092
QY 1069 GCGAFVRSQTKVPEPEVNPPLIVGSSVGGLLALLALITAAALKYKLGFFKQYKMMSE 1128
Db      :      :      :      :      :      :      :      :      :      :
1093 QCEAFLEAQTATVLEMYKVNHPVPLIVGSSVGGLLALLAITAILYKAGFFKQYKEMLEE 1152

RESULT 6
QYQY7 PRELIMINARY; PRT; 1161 AA.
AC QYQY7;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Alpha D integrin.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA O'Brien M.M.; Vandervieren M.; Kilgannon P.D.; Dietsch G.,
RA Gallatin W.M.;
RT "Cloning of rat alpha D, a novel beta 2 integrin."
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF021334; AAF21241.1; -.
DR HSSP; P11215; 1BHQ.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR GO; GO:0007229; P:integrin-mediated signaling pathway; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; EG-GAP; 3.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS50234; VWA; 1.
KW Integrin.
SQ
Query Match      55.0%; Score 3236.5; DB 11; Length 1161;
Best Local Similarity 57.5%; Pred. No. 1.3e-229;
Matches 648; Conservative 163; Mismatches 303; Indels 13; Gaps 9;

QY 2 NLTENAMTFOENARGFQSVVQLQGSRRVVGAPQEIIVAAHQSGLSYQCYSTGSCPEI 61
Db      :      :      :      :      :      :      :      :      :      :
```